SEQUENCE LISTING

<110>   Harms, Jerome S
        Splitter, Gary A
        Eakle, Kurt A
        Bremel, Robert D

<120>   INDUCIBLE PROTEIN EXPRESSION SYSTEM

<130>   960296.00463

<140>   10/763,976
<141>   2004-01-23

<150>   US 60/442,103
<151>   2003-01-23

<160>   29

<170>   PatentIn version 3.3

<210>   1
<211>   480
<212>   DNA
<213>   Artificial

<220>
<223>   Engineered sequence from virus and plasmid

<220>
<221>   promoter
<222>   (87)..(432)
<223>   BLV promoter

<400>   1
gaaaccagca gcggctatcc gcgcatccat gcccccgaac tgcaggagtg gggaggcacg   60

atggccgctt tggtcgaggc ggatcctagc agaaaaataa gacttgattc ccccttaaaa   120

ttacaactgc tagaaatga atggctctcc cgcctttttt gaggggggaat catttgtatg   180

aaagatcatg ccgacctagg cgccgccacc gccccgtaaa ccagacagag acgtcagctg   240

ccagaaaagc tggtgacggc agctggtggc tagaatcccc gtacctcccc aacttcccct   300

ttcccgaaaa atccacaccc tgagctgctg acctcacctg ctgataaatt aataaaatgc   360

cggccctgtc gagttagcgg caccagaagc gttcttctcc tgagaccctc gtgctcagct   420

ctcggtcctg cctcgagaag cttgttatca caagtttgta caaaaaagct gaacgagaaa   480

<210>   2
<211>   929
<212>   DNA
<213>   Bovine leukemia virus

1

<220>
<221> CDS
<222> (1)..(927)

<400> 2
```
atg gca agt gtt gtt ggt tgg ggg ccc cac tct cta cat gcc tgc ccg        48
Met Ala Ser Val Val Gly Trp Gly Pro His Ser Leu His Ala Cys Pro
1               5                   10                  15

gcc ctg gtt ttg tcc aat gac gtc acc atc gat gcc tgg tgc ccc ctc        96
Ala Leu Val Leu Ser Asn Asp Val Thr Ile Asp Ala Trp Cys Pro Leu
                20                  25                  30

tgc ggg ccc cat gag cga ctc caa ttc gaa agg atc gac acc acg cac       144
Cys Gly Pro His Glu Arg Leu Gln Phe Glu Arg Ile Asp Thr Thr His
            35                  40                  45

acc tgc gag acc cac cgt atc acc tgg acc gcc gat gga cga cct ttc       192
Thr Cys Glu Thr His Arg Ile Thr Trp Thr Ala Asp Gly Arg Pro Phe
        50                  55                  60

ggc ctc aat gga gcg ctg ttc cct cga ctg cat gtc tcc aga gac ccg       240
Gly Leu Asn Gly Ala Leu Phe Pro Arg Leu His Val Ser Arg Asp Pro
65                  70                  75                  80

gcc cca agg gcc cga cga ctc tgg atc aac tgc ccc ctt ccg gcc gtt       288
Ala Pro Arg Ala Arg Arg Leu Trp Ile Asn Cys Pro Leu Pro Ala Val
                85                  90                  95

cgc gct cag ccc ggc ccg gtt tca ctt tcc ccc ttc gag cgg tcc ccc       336
Arg Ala Gln Pro Gly Pro Val Ser Leu Ser Pro Phe Glu Arg Ser Pro
            100                 105                 110

ttc cag ccc tac caa tgc caa ttg ccc tcg gcc tct agc gac ggt tgc       384
Phe Gln Pro Tyr Gln Cys Gln Leu Pro Ser Ala Ser Ser Asp Gly Cys
            115                 120                 125

ccc gtc atc ggg cac ggc ctt ctt ccc tgg aac aac tta gta acg cat       432
Pro Val Ile Gly His Gly Leu Leu Pro Trp Asn Asn Leu Val Thr His
        130                 135                 140

cct tgt cct cgg aaa gtc ctt ata tta aat caa atg gcc aat ttt tcc       480
Pro Cys Pro Arg Lys Val Leu Ile Leu Asn Gln Met Ala Asn Phe Ser
145                 150                 155                 160

tta ctc ccc ccc ttc aat acc ctc ctt gtg gac ccc ctc cgg ttg tcc       528
Leu Leu Pro Pro Phe Asn Thr Leu Leu Val Asp Pro Leu Arg Leu Ser
                165                 170                 175

gtc ttt gcc cca gac acc agg gga gcc ata cgt tat ctc tcc acc ctt       576
Val Phe Ala Pro Asp Thr Arg Gly Ala Ile Arg Tyr Leu Ser Thr Leu
            180                 185                 190

ttg acg cta tgc cca gct act tgt att cta ccc ctc ggc gag ccc ttc       624
Leu Thr Leu Cys Pro Ala Thr Cys Ile Leu Pro Leu Gly Glu Pro Phe
```

2

```
tct cct aat gtc ccc ata tgt cgc ttt ccc cgg gac tcc aat gaa ccc     672
Ser Pro Asn Val Pro Ile Cys Arg Phe Pro Arg Asp Ser Asn Glu Pro
    210                 215                 220

ccc ctt tca gaa ttc gag ctg ccc ctt atc caa acg ccc ggc ctg tct     720
Pro Leu Ser Glu Phe Glu Leu Pro Leu Ile Gln Thr Pro Gly Leu Ser
225                 230                 235                 240

tgg tct gtc ccc gcg atc gac cta ttc cta acc ggc ccc cct tcc cca     768
Trp Ser Val Pro Ala Ile Asp Leu Phe Leu Thr Gly Pro Pro Ser Pro
                245                 250                 255

tgc gac cgg tta cac gta tgg tcc agt cct cag gcc tta cag cgc ttc     816
Cys Asp Arg Leu His Val Trp Ser Ser Pro Gln Ala Leu Gln Arg Phe
                260                 265                 270

ctc cat gac cct acg cta acc tgg tca gaa ttg gtt gct agc agg aaa     864
Leu His Asp Pro Thr Leu Thr Trp Ser Glu Leu Val Ala Ser Arg Lys
                275                 280                 285

cta aga ctt gat tca ccc tta aaa tta caa ctg tta gaa aat gaa tgg     912
Leu Arg Leu Asp Ser Pro Leu Lys Leu Gln Leu Leu Glu Asn Glu Trp
    290                 295                 300

ctc tcc cgc ctt ttt tg                                              929
Leu Ser Arg Leu Phe
305
```

```
<210>   3
<211>   309
<212>   PRT
<213>   Bovine leukemia virus

<400>   3
```

```
Met Ala Ser Val Val Gly Trp Gly Pro His Ser Leu His Ala Cys Pro
1               5               10              15


Ala Leu Val Leu Ser Asn Asp Val Thr Ile Asp Ala Trp Cys Pro Leu
            20              25              30


Cys Gly Pro His Glu Arg Leu Gln Phe Glu Arg Ile Asp Thr Thr His
        35              40              45


Thr Cys Glu Thr His Arg Ile Thr Trp Thr Ala Asp Gly Arg Pro Phe
    50              55              60


Gly Leu Asn Gly Ala Leu Phe Pro Arg Leu His Val Ser Arg Asp Pro
65              70              75              80
```

```
Ala Pro Arg Ala Arg Arg Leu Trp Ile Asn Cys Pro Leu Pro Ala Val
            85                    90              95

Arg Ala Gln Pro Gly Pro Val Ser Leu Ser Pro Phe Glu Arg Ser Pro
            100               105               110

Phe Gln Pro Tyr Gln Cys Gln Leu Pro Ser Ala Ser Ser Asp Gly Cys
            115               120               125

Pro Val Ile Gly His Gly Leu Leu Pro Trp Asn Asn Leu Val Thr His
            130               135               140

Pro Cys Pro Arg Lys Val Leu Ile Leu Asn Gln Met Ala Asn Phe Ser
145               150               155               160

Leu Leu Pro Pro Phe Asn Thr Leu Leu Val Asp Pro Leu Arg Leu Ser
                165               170               175

Val Phe Ala Pro Asp Thr Arg Gly Ala Ile Arg Tyr Leu Ser Thr Leu
                180               185               190

Leu Thr Leu Cys Pro Ala Thr Cys Ile Leu Pro Leu Gly Glu Pro Phe
            195               200               205

Ser Pro Asn Val Pro Ile Cys Arg Phe Pro Arg Asp Ser Asn Glu Pro
    210               215               220

Pro Leu Ser Glu Phe Glu Leu Pro Leu Ile Gln Thr Pro Gly Leu Ser
225               230               235               240

Trp Ser Val Pro Ala Ile Asp Leu Phe Leu Thr Gly Pro Pro Ser Pro
                245               250               255

Cys Asp Arg Leu His Val Trp Ser Ser Pro Gln Ala Leu Gln Arg Phe
            260               265               270

Leu His Asp Pro Thr Leu Thr Trp Ser Glu Leu Val Ala Ser Arg Lys
            275               280               285

Leu Arg Leu Asp Ser Pro Leu Lys Leu Gln Leu Leu Glu Asn Glu Trp
    290               295               300
```

4

Leu Ser Arg Leu Phe
305

<210> 4
<211> 1062
<212> DNA
<213> Human T-cell lymphotropic virus type 1

<220>
<221> CDS
<222> (1)..(1062)

<220>
<221> promoter
<222> (1)..(353)

<400> 4

```
atg gcc cac ttc cca ggg ttt gga cag agt ctt ctt ttc gga tac cca        48
Met Ala His Phe Pro Gly Phe Gly Gln Ser Leu Leu Phe Gly Tyr Pro
1               5                   10                  15

gtc tac gtg ttt gga gac ggc gac tgg tgc ccc tgt gta caa atc tct        96
Val Tyr Val Phe Gly Asp Gly Asp Trp Cys Pro Cys Val Gln Ile Ser
                20                  25                  30

ggg gga cta tgt tcg gcc cgc cta cat cgt cac gcc cta ctg gcc acc       144
Gly Gly Leu Cys Ser Ala Arg Leu His Arg His Ala Leu Leu Ala Thr
            35                  40                  45

tgt cca gag cat cag atc acc tgg gac ccc atc gat gga cgc gtt atc       192
Cys Pro Glu His Gln Ile Thr Trp Asp Pro Ile Asp Gly Arg Val Ile
        50                  55                  60

ggc tca gct cta cag ttc ctt atc cct cga ctc ccc tcc ttc ccc acc       240
Gly Ser Ala Leu Gln Phe Leu Ile Pro Arg Leu Pro Ser Phe Pro Thr
65                  70                  75                  80

cag aga acc tct aag acc ctc aag gtc ctt acc ccg cca atc act cat       288
Gln Arg Thr Ser Lys Thr Leu Lys Val Leu Thr Pro Pro Ile Thr His
                85                  90                  95

aca acc ccc aac att cca ccc tcc ttc ctc cag gcc atg cgc aaa tac       336
Thr Thr Pro Asn Ile Pro Pro Ser Phe Leu Gln Ala Met Arg Lys Tyr
                100                 105                 110

tcc ccc ttc cga aat gga tac atg gaa ccc acc ctt ggg cag cac ctc       384
Ser Pro Phe Arg Asn Gly Tyr Met Glu Pro Thr Leu Gly Gln His Leu
        115                 120                 125

cca acc ctg tct ttt cca gac ccc gga ctc cgg ccc caa aac ctg tac       432
Pro Thr Leu Ser Phe Pro Asp Pro Gly Leu Arg Pro Gln Asn Leu Tyr
        130                 135                 140

acc ctc tgg gga ggc tcc gtt gtc tgc atg tac ctc tac cag ctt tcc       480
Thr Leu Trp Gly Gly Ser Val Val Cys Met Tyr Leu Tyr Gln Leu Ser
```

5

```
ccc ccc atc acc tgg ccc ctc ctg ccc cac gtg att ttt tgc cac ccc    528
Pro Pro Ile Thr Trp Pro Leu Leu Pro His Val Ile Phe Cys His Pro
                165                 170                 175

ggc cag ctc ggg gcc ttc ctc acc aat gtt ccc tac aag cga ata gaa    576
Gly Gln Leu Gly Ala Phe Leu Thr Asn Val Pro Tyr Lys Arg Ile Glu
                180                 185                 190

gaa ctc ctc tat aaa att tcc ctt acc aca ggg gcc cta ata att cta    624
Glu Leu Leu Tyr Lys Ile Ser Leu Thr Thr Gly Ala Leu Ile Ile Leu
            195                 200                 205

ccc gaa gac tgt ttg ccc acc acc ctt ttc cag cct gtt agg gca ccc    672
Pro Glu Asp Cys Leu Pro Thr Thr Leu Phe Gln Pro Val Arg Ala Pro
        210                 215                 220

gtc acg cta aca gcc tgg caa aac ggc ctc ctt ccg ttc cac tca acc    720
Val Thr Leu Thr Ala Trp Gln Asn Gly Leu Leu Pro Phe His Ser Thr
225                 230                 235                 240

ctc acc act cca ggc ctt att tgg aca ttt acc gat ggc acg cct atg    768
Leu Thr Thr Pro Gly Leu Ile Trp Thr Phe Thr Asp Gly Thr Pro Met
                245                 250                 255

att tcc ggg ccc tgc cct aaa gat ggc cag cca tct tta gta cta cag    816
Ile Ser Gly Pro Cys Pro Lys Asp Gly Gln Pro Ser Leu Val Leu Gln
                260                 265                 270

tcc tcc tcc ttt ata ttt cac aaa ttt caa acc aag gcc tac cac ccc    864
Ser Ser Ser Phe Ile Phe His Lys Phe Gln Thr Lys Ala Tyr His Pro
            275                 280                 285

tca ttt cta ctc tca cac ggc ctc ata cag tac tct tcc ttt cat aat    912
Ser Phe Leu Leu Ser His Gly Leu Ile Gln Tyr Ser Ser Phe His Asn
        290                 295                 300

tta cat ctc ctg ttt gaa gaa tac acc aac atc ccc att tct cta ctt    960
Leu His Leu Leu Phe Glu Glu Tyr Thr Asn Ile Pro Ile Ser Leu Leu
305                 310                 315                 320

ttt aac aaa aaa gag gca gat gac aat gac cat gag ccc caa ata tcc   1008
Phe Asn Lys Lys Glu Ala Asp Asp Asn Asp His Glu Pro Gln Ile Ser
                325                 330                 335

ccc ggg ggc tta gag cct ccc agt gaa aaa cat ttc cgc gaa aca gaa   1056
Pro Gly Gly Leu Glu Pro Pro Ser Glu Lys His Phe Arg Glu Thr Glu
                340                 345                 350

gtc tga                                                            1062
Val
```

<210>   5
<211>   353

6

<212> PRT
<213> Human T-cell lymphotropic virus type 1

<400>   5

Met Ala His Phe Pro Gly Phe Gly Gln Ser Leu Leu Phe Gly Tyr Pro
1   5   10   15

Val Tyr Val Phe Gly Asp Gly Asp Trp Cys Pro Cys Val Gln Ile Ser
            20   25   30

Gly Gly Leu Cys Ser Ala Arg Leu His Arg His Ala Leu Leu Ala Thr
        35   40   45

Cys Pro Glu His Gln Ile Thr Trp Asp Pro Ile Asp Gly Arg Val Ile
    50   55   60

Gly Ser Ala Leu Gln Phe Leu Ile Pro Arg Leu Pro Ser Phe Pro Thr
65   70   75   80

Gln Arg Thr Ser Lys Thr Leu Lys Val Leu Thr Pro Pro Ile Thr His
                85   90   95

Thr Thr Pro Asn Ile Pro Pro Ser Phe Leu Gln Ala Met Arg Lys Tyr
                100   105   110

Ser Pro Phe Arg Asn Gly Tyr Met Glu Pro Thr Leu Gly Gln His Leu
    115   120   125

Pro Thr Leu Ser Phe Pro Asp Pro Gly Leu Arg Pro Gln Asn Leu Tyr
    130   135   140

Thr Leu Trp Gly Gly Ser Val Val Cys Met Tyr Leu Tyr Gln Leu Ser
145   150   155   160

Pro Pro Ile Thr Trp Pro Leu Leu Pro His Val Ile Phe Cys His Pro
                165   170   175

Gly Gln Leu Gly Ala Phe Leu Thr Asn Val Pro Tyr Lys Arg Ile Glu
            180   185   190

Glu Leu Leu Tyr Lys Ile Ser Leu Thr Thr Gly Ala Leu Ile Ile Leu
        195   200   205

7

```
Pro Glu Asp Cys Leu Pro Thr Thr Leu Phe Gln Pro Val Arg Ala Pro
    210                 215                 220

Val Thr Leu Thr Ala Trp Gln Asn Gly Leu Leu Pro Phe His Ser Thr
225                 230                 235                 240

Leu Thr Thr Pro Gly Leu Ile Trp Thr Phe Thr Asp Gly Thr Pro Met
                245                 250                 255

Ile Ser Gly Pro Cys Pro Lys Asp Gly Gln Pro Ser Leu Val Leu Gln
                260                 265                 270

Ser Ser Ser Phe Ile Phe His Lys Phe Gln Thr Lys Ala Tyr His Pro
        275                 280                 285

Ser Phe Leu Leu Ser His Gly Leu Ile Gln Tyr Ser Ser Phe His Asn
    290                 295                 300

Leu His Leu Leu Phe Glu Glu Tyr Thr Asn Ile Pro Ile Ser Leu Leu
305                 310                 315                 320

Phe Asn Lys Lys Glu Ala Asp Asp Asn Asp His Glu Pro Gln Ile Ser
                325                 330                 335

Pro Gly Gly Leu Glu Pro Pro Ser Glu Lys His Phe Arg Glu Thr Glu
            340                 345                 350

Val
```

```
<210>   6
<211>   353
<212>   DNA
<213>   Human T-cell lymphotropic virus type 1

<400>   6
tgacaatgac catgagcccc aaatatcccc cgggggctta gagcctctca gtgaaaaaca      60

tttccgtgaa acagaagtct gagaaggtca gggcccagaa taaggctctg acgtctcccc     120

ccggaggaca gctcagcacc agctcaggct aggccctgac gtgtcccct aaagacaaat      180

cataagctca gacctccggg aagccaccgg gaaccaccca ttcctcccc atgtttgtca      240

agccgtcctc aggcgttgac gacaacccct cacctcaaaa aactttcat ggcacgcata      300

cggctcaata aaataacagg agtctataaa agcgtgggga cagttcagga ggg           353
```

8

<210> 7
<211> 456
<212> DNA
<213> Human immunodeficiency virus type 1

<400> 7

```
ctggaagggc taatttggtc ccaaagaaga caagagatcc ttgatctgtg gatctaccac      60

acacaaggct acttccctga ttggcagaat tacacaccag ggccagggat cagatatcca     120

ctgacctttg gatggtgctt caagctagta ccagttgagc cagagaaggt agaagaggcc     180

aatgaaggag agaacaacag cttgttacac cctatgagcc tgcatgggat ggaggacgcg     240

gagaaagaag tgttagtgtg gaggtttgac agcaaactag catttcatca catggcccga     300

gagctgcatc cggagtacta caaagactgc tgacatcgag ctttctacaa gggactttcc     360

gctggggact tccagggag gcgtggcctg ggcgggactg gggagtggcg tccctcagat      420

gctgcatata agcagctgct ttttgcctgt actggg                               456
```

<210> 8
<211> 306
<212> DNA
<213> Human immunodeficiency virus type 1

<220>
<221> CDS
<222> (1)..(303)

<400> 8

```
atg gag cca gta gat cct aat cta gag ccc tgg aag cat cca gga agt      48
Met Glu Pro Val Asp Pro Asn Leu Glu Pro Trp Lys His Pro Gly Ser
1               5                   10                  15

cag cct agg act gct tgt aac aat tgc tat tgt aaa aag tgt tgc ttt      96
Gln Pro Arg Thr Ala Cys Asn Asn Cys Tyr Cys Lys Lys Cys Cys Phe
            20                  25                  30

cat tgc tac gcg tgt ttc aca aga aaa ggc tta ggc atc tcc tat ggc     144
His Cys Tyr Ala Cys Phe Thr Arg Lys Gly Leu Gly Ile Ser Tyr Gly
        35                  40                  45

agg aag aag cgg aga cag cga cga aga gct cct cag gac agt cag act     192
Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala Pro Gln Asp Ser Gln Thr
    50                  55                  60

cat caa gct tct cta tca aag caa ccc gcc tcc cag tcc cga ggg gac     240
His Gln Ala Ser Leu Ser Lys Gln Pro Ala Ser Gln Ser Arg Gly Asp
65                  70                  75                  80

ccg aca ggc ccg acg gaa tcg aag aag aag gtg gag aga gag aca gag     288
```

```
Pro Thr Gly Pro Thr Glu Ser Lys Lys Lys Val Glu Arg Glu Thr Glu
            85                  90                  95

aca gat ccg ttc gat tag                                              306
Thr Asp Pro Phe Asp
            100


<210>  9
<211>  101
<212>  PRT
<213>  Human immunodeficiency virus type 1

<400>  9

Met Glu Pro Val Asp Pro Asn Leu Glu Pro Trp Lys His Pro Gly Ser
1                5                  10                  15


Gln Pro Arg Thr Ala Cys Asn Asn Cys Tyr Cys Lys Lys Cys Cys Phe
            20                  25                  30


His Cys Tyr Ala Cys Phe Thr Arg Lys Gly Leu Gly Ile Ser Tyr Gly
            35                  40                  45


Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala Pro Gln Asp Ser Gln Thr
     50                  55                  60


His Gln Ala Ser Leu Ser Lys Gln Pro Ala Ser Gln Ser Arg Gly Asp
65                  70                  75                  80


Pro Thr Gly Pro Thr Glu Ser Lys Lys Lys Val Glu Arg Glu Thr Glu
            85                  90                  95


Thr Asp Pro Phe Asp
            100


<210>  10
<211>  7685
<212>  DNA
<213>  Artificial

<220>
<223>  Engineered sequence from virus and plasmid


<220>
<221>  LTR
<222>  (149)..(737)
<223>  5' MoMuSVLTR
```

```
<220>
<221>  CDS
<222>  (1753)..(2148)
<223>  Blasticidin resistance

<220>
<221>  promoter
<222>  (2257)..(3074)
<223>  CMV IE promoter

<220>
<221>  misc_recomb
<222>  (3078)..(3102)
<223>  attB1

<220>
<221>  CDS
<222>  (3115)..(4041)
<223>  BLV Tax

<220>
<221>  misc_recomb
<222>  (4046)..(4070)
<223>  attB2

<220>
<221>  misc_signal
<222>  (4082)..(4674)
<223>  WPRE; woodchuck hepatitis virus post-transcriptional regulatory
       element

<220>
<221>  LTR
<222>  (4720)..(5313)
<223>  3' MoMuLVLTR

<400>  10
gaattaattc ataccagatc accgaaaact gtcctccaaa tgtgtccccc tcacactccc       60

aaattcgcgg gcttctgcct cttagaccac tctaccctat tccccacact caccggagcc      120

aaagccgcgg cccttccgtt tctttgcttt tgaaagaccc cacccgtagg tggcaagcta      180

gcttaagtaa cgccactttg caaggcatgg aaaaatacat aactgagaat agaaaagttc      240

agatcaaggt caggaacaaa gaaacagctg aataccaaac aggatatctg tggtaagcgg      300

ttcctgcccc ggctcagggc caagaacaga tgagacagct gagtgatggg ccaaacagga      360

tatctgtggt aagcagttcc tgccccggct cggggccaag aacagatggt ccccagatgc      420

ggtccagccc tcagcagttt ctagtgaatc atcagatgtt ccagggtgc cccaaggacc       480

tgaaaatgac cctgtacctt atttgaacta accaatcagt tcgcttctcg cttctgttcg      540

cgcgcttccg ctctccgagc tcaataaaag agcccacaac ccctcactcg gcgcgccagt      600
```

11

```
cttccgatag actgcgtcgc ccgggtaccc gtattcccaa taaagcctct tgctgtttgc    660

atccgaatcg tggtctcgct gttccttggg agggtctcct ctgagtgatt gactacccac    720

gacgggggtc tttcatttgg gggctcgtcc gggatttgga gacccctgcc cagggaccac    780

cgacccacca ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag    840

tgtctatgtt tgatgttatg cgcctgcgtc tgtactagtt agctaactag ctctgtatct    900

ggcggacccg tggtggaact gacgagttct gaacacccgg ccgcaaccct gggagacgtc    960

ccagggactt tgggggccgt ttttgtggcc cgacctgagg aagggagtcg atgtggaatc   1020

cgaccccgtc aggatatgtg gttctggtag gagacgagaa cctaaaacag ttcccgcctc   1080

cgtctgaatt tttgctttcg gtttggaacc gaagccgcgc gtcttgtctg ctgcagcgct   1140

gcagcatcgt tctgtgttgt ctctgtctga ctgtgtttct gtatttgtct gaaaattagg   1200

gccagactgt taccactccc ttaagtttga ccttaggtca ctggaaagat gtcgagcgga   1260

tcgctcacaa ccagtcggta gatgtcaaga agagacgttg ggttaccttc tgctctgcag   1320

aatggccaac ctttaacgtc ggatggccgc gagacggcac ctttaaccga gacctcatca   1380

cccaggttaa gatcaaggtc ttttcacctg gcccgcatgg acacccagac caggtcccct   1440

acatcgtgac ctgggaagcc ttggcttttg acccccctcc ctgggtcaag ccctttgtac   1500

accctaagcc tccgcctcct cttcctccat ccgccccgtc tctccccctt gaacctcctc   1560

gttcgacccc gcctcgatcc tccctttatc cagccctcac tccttctcta ggcgccggaa   1620

ttccgatctg atcaagagac aggatgaggg agcttgtata tccattttcg gatctgatca   1680

gcacgtgttg acaattaatc atcggcatag tatatcggca tagtataata cgacaaggtg   1740

aggaactaaa cc atg gcc aag cct ttg tct caa gaa gaa tcc acc ctc att   1791
                Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile
                 1           5                  10

gaa aga gca acg gct aca atc aac agc atc ccc atc tct gaa gac tac     1839
Glu Arg Ala Thr Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr
   15                  20                  25

agc gtc gcc agc gca gct ctc tct agc gac ggc cgc atc ttc act ggt     1887
Ser Val Ala Ser Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly
30                  35                  40                  45

gtc aat gta tat cat ttt act ggg gga cct tgt gca gaa ctc gtg gtg     1935
Val Asn Val Tyr His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val
                 50                  55                  60

ctg ggc act gct gct gct gcg gca gct ggc aac ctg act tgt atc gtc     1983
Leu Gly Thr Ala Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val
            65                  70                  75
```

12

```
gcg atc gga aat gag aac agg ggc atc ttg agc ccc tgc gga cgg tgt       2031
Ala Ile Gly Asn Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys
        80                      85                  90

cga cag gtg ctt ctc gat ctg cat cct ggg atc aaa gcg ata gtg aag       2079
Arg Gln Val Leu Leu Asp Leu His Pro Gly Ile Lys Ala Ile Val Lys
        95                     100                 105

gac agt gat gga cag ccg acg gca gtt ggg att cgt gaa ttg ctg ccc       2127
Asp Ser Asp Gly Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro
110             115                 120                 125

tct ggt tat gtg tgg gag ggc taagcacttc gtggccgagg agcaggactg          2178
Ser Gly Tyr Val Trp Glu Gly
                130

acacgtgcta cgagatttcg attccaccgc cgccttctat gaaaggttgg gcttcggaat     2238

cgttttccgg gacgccgatc cggccattag ccatattatt cattggttat atagcataaa     2298

tcaatattgg ctattggcca ttgcatacgt tgtatccata tcataatatg tacatttata     2358

ttggctcatg tccaacatta ccgccatgtt gacattgatt attgactagt tattaatagt     2418

aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta     2478

cggtaaatgg cccgcctggc tgaccgccca cgacccccg cccattgacg tcaataatga      2538

cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt     2598

tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta     2658

ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg     2718

actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt     2778

tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc     2838

accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat     2898

gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcatgtacgg tgggaggtct     2958

atataagcag agctcgttta gtgaaccgtc agatcgcctg gagacgccat ccacgctgtt     3018

ttgacctcca tagaagacac cgggaccgat ccagcctccg cggccccaag cttgttatca     3078

caagtttgta caaaaaagca ggctcccgcc gccacc atg gca agt gtt gtt ggt       3132
                                          Met Ala Ser Val Val Gly
                                                          135

tgg ggg ccc cac tct cta cat gcc tgc ccg gcc ctg gtt ttg tcc aat       3180
Trp Gly Pro His Ser Leu His Ala Cys Pro Ala Leu Val Leu Ser Asn
        140                 145                 150

gat gtc acc atc gat gcc tgg tgc ccc ctc tgc ggg ccc cat gag cga       3228
Asp Val Thr Ile Asp Ala Trp Cys Pro Leu Cys Gly Pro His Glu Arg
```

13

ctc caa ttc gaa agg atc gac acc acg ctc acc tgc gag acc cac cgt   3276
Leu Gln Phe Glu Arg Ile Asp Thr Thr Leu Thr Cys Glu Thr His Arg
               175   180   185

atc aac tgg acc gcc gat gga cga cct tgc ggc ctc aat gga acg ttg   3324
Ile Asn Trp Thr Ala Asp Gly Arg Pro Cys Gly Leu Asn Gly Thr Leu
               190   195   200

ttc cct cga ctg cat gtc tcc gag acc cgc ccc caa ggg ccc cga cga   3372
Phe Pro Arg Leu His Val Ser Glu Thr Arg Pro Gln Gly Pro Arg Arg
               205   210   215

ctc tgg atc aac tgc ccc ctt ccg gcc gtt cgc gct cag ccc ggc ccg   3420
Leu Trp Ile Asn Cys Pro Leu Pro Ala Val Arg Ala Gln Pro Gly Pro
               220   225   230

gtt tca ctt tcc ccc ttc gag cgg tcc ccc ttc cag ccc tac caa tgc   3468
Val Ser Leu Ser Pro Phe Glu Arg Ser Pro Phe Gln Pro Tyr Gln Cys
235   240   245   250

caa ttg ccc tcg gcc tct agc gac ggt tgc ccc att atc ggg cac ggc   3516
Gln Leu Pro Ser Ala Ser Ser Asp Gly Cys Pro Ile Ile Gly His Gly
               255   260   265

ctt ctt ccc tgg aac aac tta gta acg cat cct gtc ctc aga aaa gtc   3564
Leu Leu Pro Trp Asn Asn Leu Val Thr His Pro Val Leu Arg Lys Val
               270   275   280

ctt ata tta aat caa atg gcc aat ttt tcc tta ctc ccc tcc ttc gat   3612
Leu Ile Leu Asn Gln Met Ala Asn Phe Ser Leu Leu Pro Ser Phe Asp
               285   290   295

acc ctc ctt gtg gac ccc ctc cgg ctg tcc gtc ttt gcc cca gac acc   3660
Thr Leu Leu Val Asp Pro Leu Arg Leu Ser Val Phe Ala Pro Asp Thr
      300   305   310

agg gga gcc ata cgt tat ctc tcc acc ctt ttg acg cta tgc ccg gct   3708
Arg Gly Ala Ile Arg Tyr Leu Ser Thr Leu Leu Thr Leu Cys Pro Ala
315   320   325   330

act tgt att cta ccc cta ggc gag ccc ttc tct cct aat gtc ccc ata   3756
Thr Cys Ile Leu Pro Leu Gly Glu Pro Phe Ser Pro Asn Val Pro Ile
               335   340   345

tgc cgc ttt ccc cgg gac tcc aat gaa ccc ccc ctt tca gaa ttc gag   3804
Cys Arg Phe Pro Arg Asp Ser Asn Glu Pro Pro Leu Ser Glu Phe Glu
               350   355   360

ctg ccc ctt atc caa acg ccc ggc ctg tct tgg tct gtc ccc gcg atc   3852
Leu Pro Leu Ile Gln Thr Pro Gly Leu Ser Trp Ser Val Pro Ala Ile
               365   370   375

gac cta ttc cta acc ggt ccc cct tcc cca tgc gac cgg tta cac gta   3900
Asp Leu Phe Leu Thr Gly Pro Pro Ser Pro Cys Asp Arg Leu His Val
               380   385   390

```
tgg tcc agt cct cag gcc tta cag cgc ttc ctt cat gac cct acg cta    3948
Trp Ser Ser Pro Gln Ala Leu Gln Arg Phe Leu His Asp Pro Thr Leu
395             400                 405             410

acc tgg tcc gaa tta gtt gct agc aga aaa ata aga ctt gat tcc ccc    3996
Thr Trp Ser Glu Leu Val Ala Ser Arg Lys Ile Arg Leu Asp Ser Pro
            415                 420                 425

tta aaa tta caa ctg cta gaa aat gaa tgg ctc tcc cgc ctt ttt        4041
Leu Lys Leu Gln Leu Leu Glu Asn Glu Trp Leu Ser Arg Leu Phe
            430                 435                 440

tgagacccag ctttcttgta caaagtggtg ataacatcga taatcaacct ctggattaca    4101

aaatttgtga agattgact ggtattctta actatgttgc tccttttacg ctatgtggat       4161

acgctgcttt aatgcctttg tatcatgcta ttgcttcccg tatggctttc attttctcct    4221

ccttgtataa atcctggttg ctgtctcttt atgaggagtt gtggcccgtt gtcaggcaac    4281

gtggcgtggt gtgcactgtg tttgctgacg caacccccac tggttggggc attgccacca    4341

cctgtcagct cctttccggg actttcgctt tccccctccc tattgccacg gcggaactca    4401

tcgccgcctg ccttgcccgc tgctggacag gggctcggct gttgggcact gacaattccg    4461

tggtgttgtc ggggaaatca tcgtcctttc cttggctgct cgcctgtgtt gccacctgga    4521

ttctgcgcgg gacgtccttc tgctacgtcc cttcggccct caatccagcg gaccttcctt    4581

cccgcggcct gctgccggct ctgcggcctc ttccgcgtct tcgccttcgc cctcagacga    4641

gtcggatctc cctttgggcc gcctccccgc ctgatcgata aaataaaaga ttttatttag    4701

tctccagaaa aaggggggaa tgaaagaccc cacctgtagg tttggcaagc tagcttaagt    4761

aacgccattt tgcaaggcat ggaaaaatac ataactgaga atagagaagt tcagatcaag    4821

gtcaggaaca gatggaacag ctgaatatgg gccaaacagg atatctgtgg taagcagttc    4881

ctgccccggc tcagggccaa gaacagatgg aacagctgaa tatgggccaa acaggatatc    4941

tgtggtaagc agttcctgcc ccggctcagg gccaagaaca gatggtcccc agatgcggtc    5001

cagccctcag cagtttctag agaaccatca gatgtttcca gggtgcccca aggacctgaa    5061

atgaccctgt gccttatttg aactaaccaa tcagttcgct tctcgcttct gttcgcgcgc    5121

ttctgctccc cgagctcaat aaaagagccc acaacccctc actcggggcg ccagtcctcc    5181

gattgactga gtcgcccggg tacccgtgta tccaataaac cctcttgcag ttgcatccga    5241

cttgtggtct cgctgttcct tgggagggtc tcctctgagt gattgactac cgtcagcgg     5301

gggtctttca ttttttccatt gggggctcgt ccgggatcgg gagacccctg cccagggacc    5361
```

15

```
accgacccac caccgggagg taagctggct gcctcgcgcg tttcggtgat gacggtgaaa   5421

acctctgaca catgcagctc ccggagacgg tcacagcttg tctgtaagcg gatgccggga   5481

gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg gtgtcggggc gcagccatga   5541

cccagtcacg tagcgatagc ggagtgtata ctggcttaac tatgcggcat cagagcagat   5601

tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata   5661

ccgcatcagg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct   5721

gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga   5781

taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc   5841

cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg   5901

ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg   5961

aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt   6021

tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt   6081

gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg   6141

cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact   6201

ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt   6261

cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct   6321

gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac   6381

cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc   6441

tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg   6501

ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta   6561

aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca   6621

atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc   6681

ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc   6741

tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc   6801

agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat   6861

taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt   6921

tgccattgct gcaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc   6981

cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag   7041

ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt   7101
```

16

```
tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac    7161

tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg    7221

cccggcgtca acacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat    7281

tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc    7341

gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc    7401

tgggtgagca aaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa      7461

atgttgaata ctcatactct tcctttttca atattattga agcatttatc agggttattg    7521

tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg    7581

cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac    7641

ctataaaaat aggcgtatca cgaggccctt tcgtcttcaa gaat                     7685
```

<210> 11
<211> 132
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

<400> 11

```
Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala
1               5                   10                  15


Thr Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala
            20                  25                  30


Ser Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val
            35                  40                  45


Tyr His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr
        50                  55                  60


Ala Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly
65                  70                  75                  80


Asn Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val
                85                  90                  95


Leu Leu Asp Leu His Pro Gly Ile Lys Ala Ile Val Lys Asp Ser Asp
```

17
```

Gly Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr
        115                 120                 125

Val Trp Glu Gly
    130

<210> 12
<211> 309
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

<400> 12

Met Ala Ser Val Val Gly Trp Gly Pro His Ser Leu His Ala Cys Pro
1                 5                 10                  15

Ala Leu Val Leu Ser Asn Asp Val Thr Ile Asp Ala Trp Cys Pro Leu
            20                  25                  30

Cys Gly Pro His Glu Arg Leu Gln Phe Glu Arg Ile Asp Thr Thr Leu
        35                  40                  45

Thr Cys Glu Thr His Arg Ile Asn Trp Thr Ala Asp Gly Arg Pro Cys
    50                  55                  60

Gly Leu Asn Gly Thr Leu Phe Pro Arg Leu His Val Ser Glu Thr Arg
65                  70                  75                      80

Pro Gln Gly Pro Arg Arg Leu Trp Ile Asn Cys Pro Leu Pro Ala Val
                85                  90                  95

Arg Ala Gln Pro Gly Pro Val Ser Leu Ser Pro Phe Glu Arg Ser Pro
            100                 105                 110

Phe Gln Pro Tyr Gln Cys Gln Leu Pro Ser Ala Ser Ser Asp Gly Cys
        115                 120                 125

Pro Ile Ile Gly His Gly Leu Leu Pro Trp Asn Asn Leu Val Thr His
    130                 135                 140

```
Pro Val Leu Arg Lys Val Leu Ile Leu Asn Gln Met Ala Asn Phe Ser
145             150             155             160


Leu Leu Pro Ser Phe Asp Thr Leu Leu Val Asp Pro Leu Arg Leu Ser
            165             170             175


Val Phe Ala Pro Asp Thr Arg Gly Ala Ile Arg Tyr Leu Ser Thr Leu
            180             185             190


Leu Thr Leu Cys Pro Ala Thr Cys Ile Leu Pro Leu Gly Glu Pro Phe
        195             200             205


Ser Pro Asn Val Pro Ile Cys Arg Phe Pro Arg Asp Ser Asn Glu Pro
    210             215             220


Pro Leu Ser Glu Phe Glu Leu Pro Leu Ile Gln Thr Pro Gly Leu Ser
225             230             235             240


Trp Ser Val Pro Ala Ile Asp Leu Phe Leu Thr Gly Pro Pro Ser Pro
            245             250             255


Cys Asp Arg Leu His Val Trp Ser Ser Pro Gln Ala Leu Gln Arg Phe
        260             265             270


Leu His Asp Pro Thr Leu Thr Trp Ser Glu Leu Val Ala Ser Arg Lys
    275             280             285


Ile Arg Leu Asp Ser Pro Leu Lys Leu Gln Leu Leu Glu Asn Glu Trp
    290             295             300


Leu Ser Arg Leu Phe
305



<210>   13
<211>   7685
<212>   DNA
<213>   Artificial

<220>
<223>   Engineered sequence form virus and plasmid


<220>
<221>   CDS
<222>   (210)..(1070)
<223>   Ampicillin resistance
```

<400> 13

```
attcttgaag acgaaagggc ctcgtgatac gcctattttt ataggttaat gtcatgataa      60

taatggtttc ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga accccctattt     120

gtttattttt ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa     180

tgcttcaata atattgaaaa aggaagagt atg agt att caa cat ttc cgt gtc      233
                                Met Ser Ile Gln His Phe Arg Val
                                1               5

gcc ctt att ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac      281
Ala Leu Ile Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His
        10              15                  20

cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gca      329
Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala
25                  30                  35                  40

cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag      377
Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu
                45                  50                  55

agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt      425
Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val
            60                  65                  70

ctg cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg caa gag caa      473
Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly Gln Glu Gln
            75                  80                  85

ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag tac tca      521
Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser
        90                  95                  100

cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta      569
Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu
105                 110                 115                 120

tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt      617
Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu
                125                 130                 135

ctg aca acg atc gga gga ccg aag gag cta acc gct ttt ttg cac aac      665
Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn
                140                 145                 150

atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat      713
Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn
            155                 160                 165

gaa gcc ata cca aac gac gag cgt gac acc acg atg cct gca gca atg      761
Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Ala Ala Met
        170                 175                 180

gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct      809
```

20

```
Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala
185            190             195              200

tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga      857
Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly
                205             210              215

cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa      905
Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys
            220             225              230

tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg      953
Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly
        235             240              245

cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt     1001
Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser
    250             255              260

cag gca act atg gat gaa cga aat aga cag atc gct gag ata ggt gcc     1049
Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala
265             270             275              280

tca ctg att aag cat tgg taa ctgtcagacc aagtttactc atatatactt        1100
Ser Leu Ile Lys His Trp
            285

tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat ccttttttgat  1160

aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta   1220

gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa   1280

acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt    1340

tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag    1400

ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta   1460

atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca    1520

agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag    1580

cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa   1640

agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga    1700

acaggagagc gcacgaggga gcttccaggg gaaacgcct  ggtatcttta tagtcctgtc    1760

gggtttcgcc acctctgact tgagcgtcga ttttgtgat  gctcgtcagg ggggcggagc    1820

ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt   1880

gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt   1940

gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag   2000
```

21

```
gaagcggaag agcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac   2060

cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagtataca   2120

ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgcca acacccgctg   2180

acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct   2240

ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg aggcagccag   2300

cttacctccc ggtggtgggt cggtggtccc tgggcagggg tctcccgatc ccggacgagc   2360

ccccaatgga aaaatgaaag accccgctg acgggtagtc aatcactcag aggagaccct   2420

cccaaggaac agcgagacca caagtcggat gcaactgcaa gagggtttat tggatacacg   2480

ggtacccggg cgactcagtc aatcggagga ctggcgcccc gagtgagggg ttgtgggctc   2540

ttttattgag ctcggggagc agaagcgcgc gaacagaagc gagaagcgaa ctgattggtt   2600

agttcaaata aggcacaggg tcatttcagg tccttggggc accctggaaa catctgatgg   2660

ttctctagaa actgctgagg gctggaccgc atctggggac catctgttct tggccctgag   2720

ccggggcagg aactgcttac cacagatatc ctgtttggcc catattcagc tgttccatct   2780

gttcttggcc ctgagccggg gcaggaactg cttaccacag atatcctgtt tggcccatat   2840

tcagctgttc catctgttcc tgaccttgat ctgaacttct ctattctcag ttatgtattt   2900

ttccatgcct tgcaaaatgg cgttacttaa gctagcttgc caaacctaca ggtggggtct   2960

ttcattcccc ccttttctg gagactaaat aaaatctttt attttatcga tcaggcgggg   3020

aggcggccca aagggagatc cgactcgtct gagggcgaag gcgaagacgc ggaagaggcc   3080

gcagagccgg cagcaggccg cgggaaggaa ggtccgctgg attgagggcc gaagggacgt   3140

agcagaagga cgtcccgcgc agaatccagg tggcaacaca ggcgagcagc caaggaaagg   3200

acgatgattt ccccgacaac accacggaat tgtcagtgcc caacagccga gccctgtcc    3260

agcagcgggc aaggcaggcg gcgatgagtt ccgccgtggc aatagggagg gggaaagcga   3320

aagtcccgga aaggagctga caggtggtgg caatgcccca accagtgggg gttgcgtcag   3380

caaacacagt gcacaccacg ccacgttgcc tgacaacggg ccacaactcc tcataaagag   3440

acagcaacca ggatttatac aaggaggaga aaatgaaagc catacgggaa gcaatagcat   3500

gatacaaagg cattaaagca gcgtatccac atagcgtaaa aggagcaaca tagttaagaa   3560

taccagtcaa tctttcacaa attttgtaat ccagaggttg attatcgatg ttatcaccac   3620

tttgtacaag aaagctgggt ctcaaaaaag gcgggagagc cattcatttt ctagcagttg   3680

taattttaag ggggaatcaa gtcttatttt tctgctagca actaattcgg accaggttag   3740
```

22

```
cgtagggtca tgaaggaagc gctgtaaggc ctgaggactg gaccatacgt gtaaccggtc    3800

gcatggggaa gggggaccgg ttaggaatag gtcgatcgcg gggacagacc aagacaggcc    3860

gggcgtttgg ataaggggca gctcgaattc tgaaaggggg ggttcattgg agtcccgggg    3920

aaagcggcat atggggacat taggagagaa gggctcgcct aggggtagaa tacaagtagc    3980

cgggcatagc gtcaaaaggg tggagagata acgtatggct cccctggtgt ctggggcaaa    4040

gacggacagc cggagggggt ccacaaggag ggtatcgaag gaggggagta aggaaaaatt    4100

ggccatttga tttaatataa ggacttttct gaggacagga tgcgttacta agttgttcca    4160

gggaagaagg ccgtgcccga taatggggca accgtcgcta gaggccgagg gcaattggca    4220

ttggtagggc tggaaggggg accgctcgaa gggggaaagt gaaaccgggc cgggctgagc    4280

gcgaacggcc ggaagggggc agttgatcca gagtcgtcgg ggcccttggg ggcgggtctc    4340

ggagacatgc agtcgaggga acaacgttcc attgaggccg caaggtcgtc catcggcggt    4400

ccagttgata cggtgggtct cgcaggtgag cgtggtgtcg atcctttcga attggagtcg    4460

ctcatggggc ccgcagaggg ggcaccaggc atcgatggtg acatcattgg acaaaaccag    4520

ggccgggcag gcatgtagag agtggggccc ccaaccaaca acacttgcca tggtggcggc    4580

gggagcctgc tttttgtac aaacttgtga taacaagctt ggggccgcgg aggctggatc     4640

ggtcccggtg tcttctatgg aggtcaaaac agcgtggatg gcgtctccag gcgatctgac    4700

ggttcactaa acgagctctg cttatataga cctcccaccg tacatgccta ccgcccattt    4760

gcgtcaatgg ggcggagttg ttacgacatt ttggaaagtc ccgttgattt tggtgccaaa    4820

acaaactccc attgacgtca atggggtgga gacttggaaa tccccgtgag tcaaaccgct    4880

atccacgccc attgatgtac tgccaaaacc gcatcaccat ggtaatagcg atgactaata    4940

cgtagatgta ctgccaagta ggaaagtccc ataaggtcat gtactgggca taatgccagg    5000

cgggccattt accgtcattg acgtcaatag ggggcgtact tggcatatga tacacttgat    5060

gtactgccaa gtgggcagtt taccgtaaat actccaccca ttgacgtcaa tggaaagtcc    5120

ctattggcgt tactatggga acatacgtca ttattgacgt caatgggcgg gggtcgttgg    5180

gcggtcagcc aggcgggcca tttaccgtaa gttatgtaac gcggaactcc atatatgggc    5240

tatgaactaa tgaccccgta attgattact attaataact agtcaataat caatgtcaac    5300

atggcggtaa tgttggacat gagccaatat aaatgtacat attatgatat ggatacaacg    5360

tatgcaatgg ccaatagcca atattgattt atgctatata accaatgaat aatatggcta    5420
```

23

```
atggccggat cggcgtcccg gaaaacgatt ccgaagccca acctttcata gaaggcggcg   5480

gtggaatcga aatctcgtag cacgtgtcag tcctgctcct cggccacgaa gtgcttagcc   5540

ctcccacaca taaccagagg gcagcaattc acgaatccca actgccgtcg gctgtccatc   5600

actgtccttc actatcgctt tgatcccagg atgcagatcg agaagcacct gtcgacaccg   5660

tccgcagggg ctcaagatgc ccctgttctc atttccgatc gcgacgatac aagtcaggtt   5720

gccagctgcc gcagcagcag cagtgcccag caccacgagt tctgcacaag gtcccccagt   5780

aaaatgatat acattgacac cagtgaagat gcggccgtcg ctagagagag ctgcgctggc   5840

gacgctgtag tcttcagaga tggggatgct gttgattgta gccgttgctc tttcaatgag   5900

ggtggattct tcttgagaca aaggcttggc catggtttag ttcctcacct tgtcgtatta   5960

tactatgccg atatactatg ccgatgatta attgtcaaca cgtgctgatc agatccgaaa   6020

atggatatac aagctccctc atcctgtctc ttgatcagat cggaattccg gcgcctagag   6080

aaggagtgag ggctggataa agggaggatc gaggcggggt cgaacgagga ggttcaaggg   6140

ggagagacgg ggcggatgga ggaagaggag gcggaggctt agggtgtaca aagggcttga   6200

cccagggagg ggggtcaaaa gccaaggctt cccaggtcac gatgtagggg acctggtctg   6260

ggtgtccatg cgggccaggt gaaaagacct tgatcttaac ctgggtgatg aggtctcggt   6320

taaaggtgcc gtctcgcggc catccgacgt taaaggttgg ccattctgca gagcagaagg   6380

taacccaacg tctcttcttg acatctaccg actggttgtg agcgatccgc tcgacatctt   6440

tccagtgacc taaggtcaaa cttaagggag tggtaacagt ctggccctaa ttttcagaca   6500

aatacagaaa cacagtcaga cagagacaac acagaacgat gctgcagcgc tgcagcagac   6560

aagacgcgcg gcttcggttc caaaccgaaa gcaaaaattc agacggaggc gggaactgtt   6620

ttaggttctc gtctcctacc agaaccacat atcctgacgg ggtcggattc cacatcgact   6680

cccttcctca ggtcgggcca caaaaacggc ccccaaagtc cctgggacgt ctcccagggt   6740

tgcggccggg tgttcagaac tcgtcagttc caccacgggt ccgccagata cagagctagt   6800

tagctaacta gtacagacgc aggcgcataa catcaaacat agacactaga caatcggaca   6860

gacacagata agttgctggc cagcttacct cccggtggtg ggtcggtggt ccctgggcag   6920

gggtctccaa atcccggacg agcccccaaa tgaaagaccc ccgtcgtggg tagtcaatca   6980

ctcagaggag accctcccaa ggaacagcga gaccacgatt cggatgcaaa cagcaagagg   7040

ctttattggg aatacgggta cccgggcgac gcagtctatc ggaagactgg cgcgccgagt   7100

gagggggttgt gggctctttt attgagctcg gagagcggaa gcgcgcgaac agaagcgaga   7160
```

24

```
agcgaactga ttggttagtt caaataaggt acagggtcat tttcaggtcc ttggggcacc    7220

ctggaaacat ctgatgattc actagaaact gctgagggct ggaccgcatc tggggaccat    7280

ctgttcttgg ccccgagccg gggcaggaac tgcttaccac agatatcctg tttggcccat    7340

cactcagctg tctcatctgt tcttggccct gagccggggc aggaaccgct taccacagat    7400

atcctgtttg gtattcagct gtttctttgt tcctgacctt gatctgaact tttctattct    7460

cagttatgta tttttccatg ccttgcaaag tggcgttact taagctagct tgccacctac    7520

gggtggggtc tttcaaaagc aaagaaacgg aagggccgcg gctttggctc cggtgagtgt    7580

ggggaatagg gtagagtggt ctaagaggca gaagcccgcg aatttgggag tgtgaggggg    7640

acacatttgg aggacagttt tcggtgatct ggtatgaatt aattc                    7685
```

<210> 14
<211> 286
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

<400> 14

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5                   10                  15


Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30


Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45


Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60


Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80


Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95


Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
                100                 105                 110
```

25

```
Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120             125


Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
        130                 135             140


Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150             155                 160


Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165             170             175


Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180             185             190


Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195             200             205


Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210             215             220


Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225             230             235             240


Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
            245             250             255


Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260             265             270


Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275             280             285


<210>  15
<211>  7428
<212>  DNA
<213>  Artificial

<220>
<223>  Engineered sequence from virus and plasmid


<220>
<221>  LTR
<222>  (149)..(737)
```

<223> MoMuSVLTR

<220>
<221> CDS
<222> (1660)..(2454)
<223> Neomycin resistance

<220>
<221> misc_feature
<222> (3170)..(3194)
<223> attB1

<220>
<221> CDS
<222> (3267)..(3737)
<223> trans-dominant BLV Rex (M4)

<220>
<221> misc_feature
<222> (3800)..(3824)
<223> attB2

<220>
<221> misc_signal
<222> (3837)..(4428)
<223> WPRE; woodchuck hepatitis virus post-transcriptional regulatory
      element

<220>
<221> LTR
<222> (4474)..(5067)
<223> 3' MoMuLVLTR

<400> 15
gaattaattc ataccagatc accgaaaact gtcctccaaa tgtgtccccc tcacactccc   60

aaattcgcgg gcttctgcct cttagaccac tctaccctat tccccacact caccggagcc   120

aaagccgcgg cccttccgtt tctttgcttt tgaaagaccc caccgtagg tggcaagcta   180

gcttaagtaa cgccactttg caaggcatgg aaaaatacat aactgagaat agaaaagttc   240

agatcaaggt caggaacaaa gaaacagctg aataccaaac aggatatctg tggtaagcgg   300

ttcctgcccc ggctcagggc caagaacaga tgagacagct gagtgatggg ccaaacagga   360

tatctgtggt aagcagttcc tgccccggct cggggccaag aacagatggt ccccagatgc   420

ggtccagccc tcagcagttt ctagtgaatc atcagatgtt ccagggtgc cccaaggacc   480

tgaaaatgac cctgtacctt atttgaacta accaatcagt tcgcttctcg cttctgttcg   540

cgcgcttccg ctctccgagc tcaataaaag agcccacaac ccctcactcg gcgcgccagt   600

cttccgatag actgcgtcgc ccgggtaccc gtattcccaa taaagcctct tgctgtttgc   660

atccgaatcg tggtctcgct gttccttggg agggtctcct ctgagtgatt gactacccac   720

```
gacgggggtc tttcatttgg gggctcgtcc gggatttgga gacccctgcc cagggaccac      780

cgacccacca ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag      840

tgtctatgtt tgatgttatg cgcctgcgtc tgtactagtt agctaactag ctctgtatct      900

ggcggacccg tggtggaact gacgagttct gaacacccgg ccgcaaccct gggagacgtc      960

ccagggactt tggggggccgt ttttgtggcc cgacctgagg aagggagtcg atgtggaatc     1020

cgaccccgtc aggatatgtg gttctggtag gagacgagaa cctaaaacag ttcccgcctc     1080

cgtctgaatt tttgctttcg gtttggaacc gaagccgcgc gtcttgtctg ctgcagcgct     1140

gcagcatcgt tctgtgttgt ctctgtctga ctgtgtttct gtatttgtct gaaaattagg     1200

gccagactgt taccactccc ttaagtttga ccttaggtca ctggaaagat gtcgagcgga     1260

tcgctcacaa ccagtcggta gatgtcaaga agagacgttg ggttaccttc tgctctgcag     1320

aatggccaac ctttaacgtc ggatggccgc gagacggcac ctttaaccga gacctcatca     1380

cccaggttaa gatcaaggtc ttttcacctg gcccgcatgg acacccagac caggtcccct     1440

acatcgtgac ctgggaagcc ttggcttttg accccccctcc ctgggtcaag ccctttgtac     1500

accctaagcc tccgcctcct cttcctccat ccgccccgtc tctccccctt gaacctcctc     1560

gttcgacccc gcctcgatcc tcccctttatc cagccctcac tccttctcta ggcgccggaa     1620

ttccgatctg atcaagagac aggatgagga tcgtttcgc atg att gaa caa gat       1674
                                            Met Ile Glu Gln Asp
                                             1                   5

gga ttg cac gca ggt tct ccg gcc gct tgg gtg gag agg cta ttc ggc       1722
Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly
                  10              15                  20

tat gac tgg gca caa cag aca atc ggc tgc tct gat gcc gcc gtg ttc       1770
Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe
              25                  30                  35

cgg ctg tca gcg cag ggg cgc ccg gtt ctt ttt gtc aag acc gac ctg       1818
Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu
          40                  45                  50

tcc ggt gcc ctg aat gaa ctg cag gac gag gca gcg cgg cta tcg tgg       1866
Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp
      55                  60                  65

ctg gcc acg acg ggc gtt cct tgc gca gct gtg ctc gac gtt gtc act       1914
Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val Thr
70                  75                  80                  85

gaa gcg gga agg gac tgg ctg cta ttg ggc gaa gtg ccg ggg cag gat       1962
Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu Val Pro Gly Gln Asp
```

28

```
ctc ctg tca tct cac ctt gct cct gcc gag aaa gta tcc atc atg gct      2010
Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met Ala
        105             110             115

gat gca atg cgg cgg ctg cat acg ctt gat ccg gct acc tgc cca ttc      2058
Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro Phe
        120             125             130

gac cac caa gcg aaa cat cgc atc gag cga gca cgt act cgg atg gaa      2106
Asp His Gln Ala Lys His Arg Ile Glu Arg Ala Arg Thr Arg Met Glu
    135             140             145

gcc ggt ctt gtc gat cag gat gat ctg gac gaa gag cat cag ggg ctc      2154
Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly Leu
150             155             160             165

gcg cca gcc gaa ctg ttc gcc agg ctc aag gcg cgc atg ccc gac ggc      2202
Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp Gly
            170             175             180

gag gat ctc gtc gtg acc cat ggc gat gcc tgc ttg ccg aat atc atg      2250
Glu Asp Leu Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile Met
            185             190             195

gtg gaa aat ggc cgc ttt tct gga ttc atc gac tgt ggc cgg ctg ggt      2298
Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly
        200             205             210

gtg gcg gac cgc tat cag gac ata gcg ttg gct acc cgt gat att gct      2346
Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile Ala
    215             220             225

gaa gag ctt ggc ggc gaa tgg gct gac cgc ttc ctc gtg ctt tac ggt      2394
Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr Gly
230             235             240             245

atc gcc gct ccc gat tcg cag cgc atc gcc ttc tat cgc ctt ctt gac      2442
Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp
            250             255             260

gag ttc ttc tga gcgggactct ggggttcgaa atgaccgacc aagcgacgcc          2494
Glu Phe Phe

caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg    2554

aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt    2614

cttcgcccac cccgggctcg atcccctcgc gagttggttc agctgctgcc tgaggctgga    2674

cgacctcgcg gagttctacc ggcagtgcaa atccgtcggc atccaggaaa ccagcagcgg    2734

ctatccgcgc atccatgccc ccgaactgca ggagtgggga ggcacgatgg ccgctttggt    2794

cgaggcggat cctagcagaa aaataagact tgattccccc ttaaaattac aactgctaga    2854
```

29

aaatgaatgg ctctcccgcc tttttgagg gggaatcatt tgtatgaaag atcatgccga   2914

cctaggcgcc gccaccgccc cgtaaaccag acagagacgt cagctgccag aaaagctggt   2974

gacggcagct ggtggctaga atccccgtac ctccccaact tccccttttcc cgaaaaatcc   3034

acaccctgag ctgctgacct cacctgctga taaattaata aaatgccggc cctgtcgagt   3094

tagcggcacc agaagcgttc ttctcctgag accctcgtgc tcagctctcg gtcctgcctc   3154

gagaagcttg ttatcacaag tttgtacaaa aaagcaggct cgaaggaga tagaaccaat   3214

tctctaagga aatacttaac gtcgactgga tccggtaccg aattcgatcc ac atg cct   3272
                                                                   Met Pro
                                                                   265

aaa aaa cga cgg tcc cga aga cgc cca caa ccg atc atc aga tgg caa   3320
Lys Lys Arg Arg Ser Arg Arg Arg Pro Gln Pro Ile Ile Arg Trp Gln
          270   275   280

gtg ttg ttg gtt ggg ggc ccc act ctc tac atg cct gcc cgg ccc tgg   3368
Val Leu Leu Val Gly Gly Pro Thr Leu Tyr Met Pro Ala Arg Pro Trp
          285   290   295

ttt tgt cca atg atg tca cca tcg atg cct ggt gcc ccc tct gcg ggc   3416
Phe Cys Pro Met Met Ser Pro Ser Met Pro Gly Ala Pro Ser Ala Gly
          300   305   310

ccc atg agc gac tcc aat tcg aaa gga tcg aca cca cgc tca cct gcg   3464
Pro Met Ser Asp Ser Asn Ser Lys Gly Ser Thr Pro Arg Ser Pro Ala
315   320   325   330

aga ccc acc gta tca act gga ccg ccg atg gac gac ctt gcg gcc tca   3512
Arg Pro Thr Val Ser Thr Gly Pro Pro Met Asp Asp Leu Ala Ala Ser
               335   340   345

atg gaa cgt tgt tcc ctc gac tgc atg tct ccg aga ccc gcc ccc aag   3560
Met Glu Arg Cys Ser Leu Asp Cys Met Ser Pro Arg Pro Ala Pro Lys
             350   355   360

ggc ccc gac gac tct gga tca act gcc ccc ttc cgg ccg ttc gcg ctc   3608
Gly Pro Asp Asp Ser Gly Ser Thr Ala Pro Phe Arg Pro Phe Ala Leu
          365   370   375

agc ccg gcc cgg tta gat ctt ccc cct tcg agc ggt ccc cct tcc agc   3656
Ser Pro Ala Arg Leu Asp Leu Pro Pro Ser Ser Gly Pro Pro Ser Ser
          380   385   390

cct acc aat gcc aat tgc cct cgg cct cta gcg acg gtt gcc cca tta   3704
Pro Thr Asn Ala Asn Cys Pro Arg Pro Leu Ala Thr Val Ala Pro Leu
395   400   405   410

tcg ggc acg gcc ttc ttc cct gga aca act tag taacgcatcc tgtcctcaga   3757
Ser Gly Thr Ala Phe Phe Pro Gly Thr Thr
                  415   420

```
aaagtcctta tattaaatca aatgggacct cgagatatct agacccagct ttcttgtaca   3817

aagtggtgat aacatcgata atcaacctct ggattacaaa atttgtgaaa gattgactgg   3877

tattcttaac tatgttgctc cttttacgct atgtggatac gctgctttaa tgcctttgta   3937

tcatgctatt gcttcccgta tggctttcat tttctcctcc ttgtataaat cctggttgct   3997

gtctctttat gaggagttgt ggcccgttgt caggcaacgt ggcgtggtgt gcactgtgtt   4057

tgctgacgca accccactg gttggggcat tgccaccacc tgtcagctcc tttccgggac    4117

tttcgctttc cccctccta ttgccacggc ggaactcatc gccgcctgcc ttgcccgctg     4177

ctggacaggg gctcggctgt tgggcactga caattccgtg gtgttgtcgg ggaaatcatc   4237

gtcctttcct tggctgctcg cctgtgttgc cacctggatt ctgcgcggga cgtccttctg    4297

ctacgtccct tcggccctca atccagcgga ccttccttcc cgcggcctgc tgccggctct   4357

gcggcctctt ccgcgtcttc gccttcgccc tcagacgagt cggatctccc tttgggccgc   4417

ctccccgcct gatcgataaa ataaaagatt ttatttagtc tccagaaaaa gggggaatg    4477

aaagacccca cctgtaggtt tggcaagcta gcttaagtaa cgccattttg caaggcatgg   4537

aaaaatacat aactgagaat agagaagttc agatcaaggt caggaacaga tggaacagct   4597

gaatatgggc caaacaggat atctgtggta agcagttcct gccccggctc agggccaaga   4657

acagatggaa cagctgaata tgggccaaac aggatatctg tggtaagcag ttcctgcccc    4717

ggctcagggc caagaacaga tggtccccag atgcggtcca gccctcagca gtttctagag   4777

aaccatcaga tgtttccagg gtgccccaag gacctgaaat gaccctgtgc cttatttgaa   4837

ctaaccaatc agttcgcttc tcgcttctgt tcgcgcgctt ctgctccccg agctcaataa   4897

aagagcccac aacccctcac tcggggcgcc agtcctccga ttgactgagt cgcccgggta    4957

cccgtgtatc caataaaccc tcttgcagtt gcatccgact tgtggtctcg ctgttccttg    5017

ggagggtctc ctctgagtga ttgactaccc gtcagcgggg gtctttcatt tggggggctcg   5077

tccgggatcg ggagacccct gcccagggac caccgaccca ccaccgggag gtaagctggc   5137

tgcctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg   5197

gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg   5257

ggtgttggcg ggtgtcgggg cgcagccatg acccagtcac gtagcgatag cggagtgtat   5317

actggcttaa ctatgcggca tcagagcaga ttgtactgag agtgcaccat atgcggtgtg   5377

aaataccgca cagatgcgta aggagaaaat accgcatcag gcgctcttcc gcttcctcgc   5437

tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg   5497
```

31

```
cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag    5557

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc    5617

gccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga acccgacag      5677

gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga    5737

ccctgccgct taccggatac ctgtataaag ataccaggcg tttccccctg gaagctccct    5797

cgtgcgctct cctgttccga ccctgccgct taccggatac ctgctccaag ctgggctgtg    5857

tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt    5917

ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca    5977

gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca    6037

ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag    6097

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca    6157

agcagcagat tacgcgcaga aaaaaggat ctcaagaatc ctttgatctt tgatctacgg     6217

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa    6277

aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta    6337

tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag    6397

cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga    6457

tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac    6517

cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc    6577

ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta    6637

gttcgccagt taatagtttg cgcaacgttg ttgccattgc tgcaggcatc gtggtgtcac    6697

gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat    6757

gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa    6817

gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg    6877

tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag    6937

aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aacacgggat aataccgcgc    6997

cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct    7057

caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat    7117

cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg    7177
```

32

```
ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc ttccttttc      7237

aatattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta      7297

tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg      7357

tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct      7417

ttcgtcttca a                                                           7428


<210>   16
<211>   264
<212>   PRT
<213>   Artificial

<220>
<223>   Synthetic Construct

<400>   16

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1               5                   10                  15


Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
            20                  25                  30


Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
            35                  40                  45


Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
    50                  55                  60


Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65                  70                  75                  80


Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
                85                  90                  95


Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
            100                 105                 110


Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
            115                 120                 125


Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
    130                 135                 140
```

33

```
Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145             150             155             160


Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
            165             170             175


Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
            180             185             190


Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
            195             200             205


Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
    210             215             220


Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225             230             235             240


Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
            245             250             255


Tyr Arg Leu Leu Asp Glu Phe Phe
            260


<210>   17
<211>   156
<212>   PRT
<213>   Artificial

<220>
<223>   Synthetic Construct

<400>   17

Met Pro Lys Lys Arg Arg Ser Arg Arg Arg Pro Gln Pro Ile Ile Arg
1               5               10              15


Trp Gln Val Leu Leu Val Gly Gly Pro Thr Leu Tyr Met Pro Ala Arg
            20              25              30


Pro Trp Phe Cys Pro Met Met Ser Pro Ser Met Pro Gly Ala Pro Ser
            35              40              45


Ala Gly Pro Met Ser Asp Ser Asn Ser Lys Gly Ser Thr Pro Arg Ser
    50              55              60
```

34

```
Pro Ala Arg Pro Thr Val Ser Thr Gly Pro Pro Met Asp Asp Leu Ala
65              70              75              80


Ala Ser Met Glu Arg Cys Ser Leu Asp Cys Met Ser Pro Arg Pro Ala
            85              90              95


Pro Lys Gly Pro Asp Asp Ser Gly Ser Thr Ala Pro Phe Arg Pro Phe
        100             105             110


Ala Leu Ser Pro Ala Arg Leu Asp Leu Pro Pro Ser Ser Gly Pro Pro
        115             120             125


Ser Ser Pro Thr Asn Ala Asn Cys Pro Arg Pro Leu Ala Thr Val Ala
    130             135             140


Pro Leu Ser Gly Thr Ala Phe Phe Pro Gly Thr Thr
145             150             155


<210>  18
<211>  7428
<212>  DNA
<213>  Artificial

<220>
<223>  Engineered sequence form virus and plasmid


<220>
<221>  CDS
<222>  (206)..(1066)
<223>  Ampicillin resistance

<400>  18
ttgaagacga aagggcctcg tgatacgcct atttttatag gttaatgtca tgataataat      60

ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt     120

atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct     180

tcaataatat tgaaaaagga agagt atg agt att caa cat ttc cgt gtc gcc     232
                          Met Ser Ile Gln His Phe Arg Val Ala
                          1               5

ctt att ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac cca     280
Leu Ile Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro
10              15              20              25

gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gca cga     328
Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg
                30              35              40
```

35

```
gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag agt        376
Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser
            45              50              55

ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg        424
Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu
        60              65              70

cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg caa gag caa ctc        472
Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly Gln Glu Gln Leu
        75              80              85

ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag tac tca cca        520
Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro
90              95              100             105

gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta tgc        568
Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys
                110             115             120

agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg        616
Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu
                125             130             135

aca acg atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg        664
Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met
            140             145             150

ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa        712
Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu
155             160             165

gcc ata cca aac gac gag cgt gac acc acg atg cct gca gca atg gca        760
Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Ala Ala Met Ala
170             175             180             185

aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc        808
Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser
            190             195             200

cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca        856
Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro
            205             210             215

ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct        904
Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser
            220             225             230

gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca        952
Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro
        235             240             245

gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag       1000
Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln
250             255             260             265
```

36

```
gca act atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca      1048
Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser
            270             275             280

ctg att aag cat tgg taa ctgtcagacc aagtttactc atatatactt           1096
Leu Ile Lys His Trp
            285


tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat ccttttttgat   1156

aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta    1216

gatcaaagat caaaggattc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa    1276

acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt     1336

tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag     1396

ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta    1456

atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca    1516

agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag     1576

cccagcttgg agcaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgaggga    1636

gcttccaggg ggaaacgcct ggtatcttta tacaggtatc cggtaagcgg cagggtcgga    1696

acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc    1756

gggtttcgcc acctctgact tgagcgtcga ttttgtgat gctcgtcagg ggggcggagc      1816

ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt    1876

gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt    1936

gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag     1996

gaagcggaag agcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac     2056

cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagtataca    2116

ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgcca cacccgctg      2176

acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct    2236

ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg aggcagccag    2296

cttacctccc ggtggtgggt cggtggtccc tgggcagggg tctcccgatc ccggacgagc     2356

ccccaaatga agaccccccg ctgacgggta gtcaatcact cagaggagac cctcccaagg    2416

aacagcgaga ccacaagtcg gatgcaactg caagagggtt tattggatac acgggtaccc    2476

gggcgactca gtcaatcgga ggactggcgc cccgagtgag gggttgtggg ctcttttatt     2536

gagctcgggg agcagaagcg cgcgaacaga agcgagaagc gaactgattg gttagttcaa     2596
```

37

```
ataaggcaca gggtcatttc aggtccttgg ggcaccctgg aaacatctga tggttctcta    2656

gaaactgctg agggctggac cgcatctggg gaccatctgt tcttggccct gagccggggc    2716

aggaactgct taccacagat atcctgtttg gcccatattc agctgttcca tctgttcttg    2776

gccctgagcc ggggcaggaa ctgcttacca cagatatcct gtttggccca tattcagctg    2836

ttccatctgt tcctgacctt gatctgaact tctctattct cagttatgta tttttccatg    2896

ccttgcaaaa tggcgttact taagctagct tgccaaacct acaggtgggg tctttcattc    2956

cccccttttt ctggagacta aataaaatct tttattttat cgatcaggcg gggaggcggc    3016

ccaaagggag atccgactcg tctgagggcg aaggcgaaga cgcggaagag gccgcagagc    3076

cggcagcagg ccgcgggaag gaaggtccgc tggattgagg gccgaaggga cgtagcagaa    3136

ggacgtcccg cgcagaatcc aggtggcaac acaggcgagc agccaaggaa aggacgatga    3196

tttccccgac aacaccacgg aattgtcagt gcccaacagc cgagcccctg tccagcagcg    3256

ggcaaggcag gcggcgatga gttccgccgt ggcaataggg aggggaaag cgaaagtccc     3316

ggaaaggagc tgacaggtgg tggcaatgcc ccaaccagtg ggggttgcgt cagcaaacac    3376

agtgcacacc acgccacgtt gcctgacaac gggccacaac tcctcataaa gagacagcaa    3436

ccaggattta tacaaggagg agaaaatgaa agccatacgg gaagcaatag catgatacaa    3496

aggcattaaa gcagcgtatc cacatagcgt aaaaggagca acatagttaa gaataccagt    3556

caatctttca caaattttgt aatccagagg ttgattatcg atgttatcac cactttgtac    3616

aagaaagctg ggtctagata tctcgaggtc ccatttgatt taatataagg acttttctga    3676

ggacaggatg cgttactaag ttgttccagg gaagaaggcc gtgcccgáta atggggcaac    3736

cgtcgctaga ggccgagggc aattggcatt ggtagggctg gaaggggac cgctcgaagg     3796

gggaagatct aaccgggccg ggctgagcgc gaacggccgg aaggggcag ttgatccaga     3856

gtcgtcgggg cccttggggg cgggtctcgg agacatgcag tcgagggaac aacgttccat    3916

tgaggccgca aggtcgtcca tcggcggtcc agttgatacg gtgggtctcg caggtgagcg    3976

tggtgtcgat cctttcgaat tggagtcgct catggggccc gcagagggg caccaggcat     4036

cgatggtgac atcattggac aaaaccaggg ccgggcaggc atgtagagag tggggccccc    4096

aaccaacaac acttgccatc tgatgatcgg ttgtgggcgt cttcgggacc gtcgtttttt    4156

aggcatgtgg atcgaattcg gtaccggatc cagtcgacgt taagtatttc cttagagaat    4216

tggttctatc tccttcgaag cctgcttttt tgtacaaact tgtgataaca agcttctcga    4276
```

38

```
ggcaggaccg agagctgagc acgagggtct caggagaaga acgcttctgg tgccgctaac   4336

tcgacagggc cggcatttta ttaatttatc agcaggtgag gtcagcagct cagggtgtgg   4396

atttttcggg aaaggggaag ttggggaggt acggggattc tagccaccag ctgccgtcac   4456

cagcttttct ggcagctgac gtctctgtct ggtttacggg gcggtggcgg cgcctaggtc   4516

ggcatgatct ttcatacaaa tgattccccc tcaaaaaagg cgggagagcc attcattttc   4576

tagcagttgt aattttaagg gggaatcaag tcttattttt ctgctaggat ccgcctcgac   4636

caaagcggcc atcgtgcctc cccactcctg cagttcgggg gcatggatgc gcggatagcc   4696

gctgctggtt tcctggatgc cgacggattt gcactgccgg tagaactccg cgaggtcgtc   4756

cagcctcagg cagcagctga accaactcgc gaggggatcg agcccggggt gggcgaagaa   4816

ctccagcatg agatccccgc gctggaggat catccagccg gcgtcccgga aaacgattcc   4876

gaagcccaac ctttcataga aggcggcggt ggaatcgaaa tctcgtgatg gcaggttggg   4936

cgtcgcttgg tcggtcattt cgaaccccag agtcccgctc agaagaactc gtcaagaagg   4996

cgatagaagg cgatgcgctg cgaatcggga gcggcgatac cgtaaagcac gaggaagcgg   5056

tcagcccatt cgccgccaag ctcttcagca atatcacggg tagccaacgc tatgtcctga   5116

tagcggtccg ccacacccag ccggccacag tcgatgaatc cagaaaagcg gccattttcc   5176

accatgatat tcggcaagca ggcatcgcca tgggtcacga cgagatcctc gccgtcgggc   5236

atgcgcgcct tgagcctggc gaacagttcg gctggcgcga gcccctgatg ctcttcgtcc   5296

agatcatcct gatcgacaag accggcttcc atccgagtac gtgctcgctc gatgcgatgt   5356

ttcgcttggt ggtcgaatgg gcaggtagcc ggatcaagcg tatgcagccg ccgcattgca   5416

tcagccatga tggatacttt ctcggcagga gcaaggtgag atgacaggag atcctgcccc   5476

ggcacttcgc ccaatagcag ccagtccctt cccgcttcag tgacaacgtc gagcacagct   5536

gcgcaaggaa cgcccgtcgt ggccagccac gatagccgcg ctgcctcgtc ctgcagttca   5596

ttcagggcac cggacaggtc ggtcttgaca aaaagaaccg ggcgcccctg cgctgacagc   5656

cggaacacgg cggcatcaga gcagccgatt gtctgttgtg cccagtcata gccgaatagc   5716

ctctccaccc aagcggccgg agaacctgcg tgcaatccat cttgttcaat catgcgaaac   5776

gatcctcatc ctgtctcttg atcagatcgg aattccggcg cctagagaag gagtgagggc   5836

tggataaagg gaggatcgag gcggggtcga acgaggaggt tcaaggggga gagacggggc   5896

ggatggagga agaggaggcg gaggcttagg gtgtacaaag ggcttgaccc agggaggggg   5956

gtcaaaagcc aaggcttccc aggtcacgat gtaggggacc tggtctgggt gtccatgcgg   6016
```

39

```
gccaggtgaa aagaccttga tcttaacctg ggtgatgagg tctcggttaa aggtgccgtc    6076

tcgcggccat ccgacgttaa aggttggcca ttctgcagag cagaaggtaa cccaacgtct    6136

cttcttgaca tctaccgact ggttgtgagc gatccgctcg acatctttcc agtgacctaa    6196

ggtcaaactt aagggagtgg taacagtctg gccctaattt tcagacaaat acagaaacac    6256

agtcagacag agacaacaca gaacgatgct gcagcgctgc agcagacaag acgcgcggct    6316

tcggttccaa accgaaagca aaaattcaga cggaggcggg aactgtttta ggttctcgtc    6376

tcctaccaga accacatatc ctgacggggt cggattccac atcgactccc ttcctcaggt    6436

cgggccacaa aaacggcccc caaagtccct gggacgtctc ccagggttgc ggccgggtgt    6496

tcagaactcg tcagttccac cacgggtccg ccagatacag agctagttag ctaactagta    6556

cagacgcagg cgcataacat caaacataga cactagacaa tcggacagac acagataagt    6616

tgctggccag cttacctccc ggtggtgggt cggtggtccc tgggcagggg tctccaaatc    6676

ccggacgagc ccccaaatga aagaccccg tcgtgggtag tcaatcactc agaggagacc    6736

ctcccaagga acagcgagac cacgattcgg atgcaaacag caagaggctt tattgggaat    6796

acgggtaccc gggcgacgca gtctatcgga agactggcgc gccgagtgag gggttgtggg    6856

ctcttttatt gagctcggag agcggaagcg cgcgaacaga agcgagaagc gaactgattg    6916

gttagttcaa ataaggtaca gggtcatttt caggtccttg gggcaccctg gaaacatctg    6976

atgattcact agaaactgct gagggctgga ccgcatctgg ggaccatctg ttcttggccc    7036

cgagccgggg caggaactgc ttaccacaga tatcctgttt ggcccatcac tcagctgtct    7096

catctgttct tggccctgag ccggggcagg aaccgcttac cacagatatc ctgtttggta    7156

ttcagctgtt tctttgttcc tgaccttgat ctgaactttt ctattctcag ttatgtattt    7216

ttccatgcct tgcaaagtgg cgttacttaa gctagcttgc cacctacggg tggggtcttt    7276

caaaagcaaa gaaacggaag ggccgcggct ttggctccgg tgagtgtggg gaatagggta    7336

gagtggtcta agaggcagaa gcccgcgaat ttgggagtgt gaggggggaca catttggagg    7396

acagttttcg gtgatctggt atgaattaat tc                                  7428
```

<210> 19
<211> 286
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

40

<400> 19

Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1   5   10   15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20   25   30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35   40   45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50   55   60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65   70   75   80

Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
            85   90   95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100   105   110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115   120   125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130   135   140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145   150   155   160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
            165   170   175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
        180   185   190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195   200   205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210   215   220

41

```
Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230             235                 240


Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245             250             255


Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
                260             265             270


Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
            275             280             285



<210>   20
<211>   7010
<212>   DNA
<213>   Artificial

<220>
<223>   Engineered sequence from virus and plasmid


<220>
<221>   LTR
<222>   (149)..(737)
<223>   5' MoMuSVLTR

<220>
<221>   CDS
<222>   (1660)..(2454)
<223>   Neomycin resistance

<220>
<221>   promoter
<222>   (2806)..(3150)
<223>   BLV promoter

<220>
<221>   misc_feature
<222>   (3170)..(3194)
<223>   attB1

<220>
<221>   CDS
<222>   (3236)..(3955)
<223>   EYFP; enhanced yellow fluorescent protein

<220>
<221>   misc_feature
<222>   (3980)..(4004)
<223>   attB2
```

<220>
<221> LTR
<222> (4056)..(4649)
<223> 3' MoMuSVLTR

<400> 20
```
gaattaattc ataccagatc accgaaaact gtcctccaaa tgtgtccccc tcacactccc     60

aaattcgcgg gcttctgcct cttagaccac tctaccctat tccccacact caccggagcc    120

aaagccgcgg cccttccgtt tctttgcttt tgaaagaccc cacccgtagg tggcaagcta    180

gcttaagtaa cgccactttg caaggcatgg aaaaatacat aactgagaat agaaaagttc    240

agatcaaggt caggaacaaa gaaacagctg ataccaaac aggatatctg tggtaagcgg     300

ttcctgcccc ggctcagggc caagaacaga tgagacagct gagtgatggg ccaaacagga    360

tatctgtggt aagcagttcc tgccccggct cggggccaag aacagatggt ccccagatgc    420

ggtccagccc tcagcagttt ctagtgaatc atcagatgtt ccagggtgc cccaaggacc     480

tgaaaatgac cctgtacctt atttgaacta accaatcagt tcgcttctcg cttctgttcg    540

cgcgcttccg ctctccgagc tcaataaaag agcccacaac ccctcactcg gcgcgccagt    600

cttccgatag actgcgtcgc ccgggtaccc gtattcccaa taaagcctct tgctgtttgc    660

atccgaatcg tggtctcgct gttccttggg agggtctcct ctgagtgatt gactacccac    720

gacgggggtc tttcatttgg gggctcgtcc gggatttgga gacccctgcc cagggaccac    780

cgacccacca ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag    840

tgtctatgtt tgatgttatg cgcctgcgtc tgtactagtt agctaactag ctctgtatct    900

ggcggacccg tggtggaact gacgagttct gaacacccgg ccgcaaccct gggagacgtc    960

ccagggactt tgggggccgt ttttgtggcc cgacctgagg aagggagtcg atgtggaatc   1020

cgaccccgtc aggatatgtg gttctggtag gagacgagaa cctaaaacag ttcccgcctc   1080

cgtctgaatt tttgctttcg gtttggaacc gaagccgcgc gtcttgtctg ctgcagcgct   1140

gcagcatcgt tctgtgttgt ctctgtctga ctgtgtttct gtatttgtct gaaaattagg   1200

gccagactgt taccactccc ttaagtttga ccttaggtca ctggaaagat gtcgagcgga   1260

tcgctcacaa ccagtcggta gatgtcaaga agagacgttg ggttaccttc tgctctgcag   1320

aatggccaac ctttaacgtc ggatggccgc gagacggcac ctttaaccga gacctcatca   1380

cccaggttaa gatcaaggtc ttttcacctg gcccgcatgg acacccagac caggtcccct   1440

acatcgtgac ctgggaagcc ttggcttttg accccccctcc ctgggtcaag ccctttgtac   1500

accctaagcc tccgcctcct cttcctccat ccgccccgtc tctcccccttt gaacctcctc   1560
```

43

```
gttcgacccc gcctcgatcc tccctttatc cagccctcac tccttctcta ggcgccggaa    1620

ttccgatctg atcaagagac aggatgagga tcgtttcgc atg att gaa caa gat        1674
                                          Met Ile Glu Gln Asp
                                          1                 5

gga ttg cac gca ggt tct ccg gcc gct tgg gtg gag agg cta ttc ggc      1722
Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly
                10              15              20

tat gac tgg gca caa cag aca atc ggc tgc tct gat gcc gcc gtg ttc      1770
Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe
            25              30              35

cgg ctg tca gcg cag ggg cgc ccg gtt ctt ttt gtc aag acc gac ctg      1818
Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu
        40              45              50

tcc ggt gcc ctg aat gaa ctg cag gac gag gca gcg cgg cta tcg tgg      1866
Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp
    55              60              65

ctg gcc acg acg ggc gtt cct tgc gca gct gtg ctc gac gtt gtc act      1914
Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val Thr
70              75              80              85

gaa gcg gga agg gac tgg ctg cta ttg ggc gaa gtg ccg ggg cag gat      1962
Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu Val Pro Gly Gln Asp
                90              95              100

ctc ctg tca tct cac ctt gct cct gcc gag aaa gta tcc atc atg gct      2010
Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met Ala
            105             110             115

gat gca atg cgg cgg ctg cat acg ctt gat ccg gct acc tgc cca ttc      2058
Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro Phe
        120             125             130

gac cac caa gcg aaa cat cgc atc gag cga gca cgt act cgg atg gaa      2106
Asp His Gln Ala Lys His Arg Ile Glu Arg Ala Arg Thr Arg Met Glu
    135             140             145

gcc ggt ctt gtc gat cag gat gat ctg gac gaa gag cat cag ggg ctc      2154
Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly Leu
150             155             160             165

gcg cca gcc gaa ctg ttc gcc agg ctc aag gcg cgc atg ccc gac ggc      2202
Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp Gly
                170             175             180

gag gat ctc gtc gtg acc cat ggc gat gcc tgc ttg ccg aat atc atg      2250
Glu Asp Leu Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile Met
            185             190             195

gtg gaa aat ggc cgc ttt tct gga ttc atc gac tgt ggc cgg ctg ggt      2298
Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly
```

44

```
gtg gcg gac cgc tat cag gac ata gcg ttg gct acc cgt gat att gct        2346
Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile Ala
    215                 220                 225

gaa gag ctt ggc ggc gaa tgg gct gac cgc ttc ctc gtg ctt tac ggt        2394
Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr Gly
230                 235                 240                 245

atc gcc gct ccc gat tcg cag cgc atc gcc ttc tat cgc ctt ctt gac        2442
Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp
                250                 255                 260

gag ttc ttc tga gcgggactct ggggttcgaa atgaccgacc aagcgacgcc           2494
Glu Phe Phe


caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg      2554

aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt      2614

cttcgcccac cccgggctcg atcccctcgc gagttggttc agctgctgcc tgaggctgga      2674

cgacctcgcg gagttctacc ggcagtgcaa atccgtcggc atccaggaaa ccagcagcgg      2734

ctatccgcgc atccatgccc ccgaactgca ggagtgggga ggcacgatgg ccgctttggt      2794

cgaggcggat cctagcagaa aaataagact tgattccccc ttaaaattac aactgctaga      2854

aaatgaatgg ctctcccgcc ttttttgagg gggaatcatt tgtatgaaag atcatgccga      2914

cctaggcgcc gccaccgccc cgtaaaccag acagagacgt cagctgccag aaaagctggt      2974

gacggcagct ggtggctaga atccccgtac ctccccaact tcccctttcc cgaaaaatcc      3034

acaccctgag ctgctgacct cacctgctga taaattaata aaatgccggc cctgtcgagt      3094

tagcggcacc agaagcgttc ttctcctgag accctcgtgc tcagctctcg gtcctgcctc      3154

gagaagcttg ttatcacaag tttgtacaaa aaagcaggct tcgaaggaga tagaaccaat      3214

tctctaagga aatacttaac c atg gtg agc aag ggc gag gag ctg ttc acc        3265
                        Met Val Ser Lys Gly Glu Glu Leu Phe Thr
                        265                 270

ggg gtg gtg ccc atc ctg gtc gag ctg gac ggc gac gta aac ggc cac        3313
Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His
275                 280                 285                 290

aag ttc agc gtg tcc ggc gag ggc gag ggc gat gcc acc tac ggc aag        3361
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys
                295                 300                 305

ctg acc ctg aag ttc atc tgc acc acc ggc aag ctg ccc gtg ccc tgg        3409
Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp
        310                 315                 320
```

45

```
ccc acc ctc gtg acc acc ttc ggc tac ggc ctg cag tgc ttc gcc cgc       3457
Pro Thr Leu Val Thr Thr Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg
        325                     330                 335

tac ccc gac cac atg aag cag cac gac ttc ttc aag tcc gcc atg ccc       3505
Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro
        340                     345                 350

gaa ggc tac gtc cag gag cgc acc atc ttc ttc aag gac gac ggc aac       3553
Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn
355                     360                 365                 370

tac aag acc cgc gcc gag gtg aag ttc gag ggc gac acc ctg gtg aac       3601
Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn
                375                     380                 385

cgc atc gag ctg aag ggc atc gac ttc aag gag gac ggc aac atc ctg       3649
Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu
            390                     395                 400

ggg cac aag ctg gag tac aac tac aac agc cac aac gtc tat atc atg       3697
Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met
            405                     410                 415

gcc gac aag cag aag aac ggc atc aag gtg aac ttc aag atc cgc cac       3745
Ala Asp Lys Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His
            420                     425                 430

aac atc gag gac ggc agc gtg cag ctc gcc gac cac tac cag cag aac       3793
Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn
435                     440                 445                 450

acc ccc atc ggc gac ggc ccc gtg ctg ctg ccc gac aac cac tac ctg       3841
Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu
                455                     460                 465

agc tac cag tcc gcc ctg agc aaa gac ccc aac gag aag cgc gat cac       3889
Ser Tyr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His
                470                     475                 480

atg gtc ctg ctg gag ttc gtg acc gcc gcc ggg atc act ctc ggc atg       3937
Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met
            485                     490                 495

gac gag ctg tac aag taa agcggccgca ctcgagatat ctagacccag             3985
Asp Glu Leu Tyr Lys
        500

ctttcttgta caaagtggtg ataacatcga taaaataaaa gattttattt agtctccaga    4045

aaaagggggg aatgaaagac cccacctgta ggtttggcaa gctagcttaa gtaacgccat    4105

tttgcaaggc atggaaaaat acataactga gaatagagaa gttcagatca aggtcaggaa    4165

cagatggaac agctgaatat gggccaaaca ggatatctgt ggtaagcagt tcctgccccg    4225
```

46

```
gctcagggcc aagaacagat ggaacagctg aatatgggcc aaacaggata tctgtggtaa    4285

gcagttcctg ccccggctca gggccaagaa cagatggtcc ccagatgcgg tccagccctc    4345

agcagtttct agagaaccat cagatgtttc cagggtgccc caaggacctg aaatgaccct    4405

gtgccttatt tgaactaacc aatcagttcg cttctcgctt ctgttcgcgc gcttctgctc    4465

cccgagctca ataaaagagc ccacaacccc tcactcgggg cgccagtcct ccgattgact    4525

gagtcgcccg ggtacccgtg tatccaataa accctcttgc agttgcatcc gacttgtggt    4585

ctcgctgttc cttgggaggg tctcctctga gtgattgact acccgtcagc gggggtcttt    4645

catttggggg ctcgtccggg atcgggagac ccctgcccag ggaccaccga cccaccaccg    4705

ggaggtaagc tggctgcctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc    4765

agctcccgga cacggtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc    4825

agggcgcgtc agcgggtgtt ggcgggtgtc ggggcgcagc catgacccag tcacgtagcg    4885

atagcggagt gtatactggc ttaactatgc ggcatcagag cagattgtac tgagagtgca    4945

ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggcgctc    5005

ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc    5065

agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa    5125

catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt    5185

tttccatagg ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg    5245

gcgaaacccg acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg    5305

ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag    5365

cgtggcgctt tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc    5425

caagctgggc tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct tatccggtaa    5485

ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg    5545

taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc    5605

taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac    5665

cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg    5725

tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt    5785

gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt    5845

catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa    5905

atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga    5965
```

47

```
ggcacctatc tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt    6025

gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg    6085

agacccacgc tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga    6145

gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga    6205

agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctgcagg    6265

catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc    6325

aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc    6385

gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca    6445

taattctctt actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac    6505

caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaacacg    6565

ggataatacc gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc    6625

ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg    6685

tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac    6745

aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat    6805

actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata    6865

catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa    6925

agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg    6985

tatcacgagg ccctttcgtc ttcaa                                          7010


<210>   21
<211>   264
<212>   PRT
<213>   Artificial

<220>
<223>   Synthetic Construct

<400>   21

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1               5                   10                  15


Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
                20                  25                  30


Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
```

48

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
    50   55   60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65   70   75   80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
                85   90   95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
            100   105   110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
        115   120   125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
    130   135   140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145   150   155   160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
            165   170   175

Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
            180   185   190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
        195   200   205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
    210   215   220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225   230   235   240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
            245   250   255

Tyr Arg Leu Leu Asp Glu Phe Phe
        260

49

<210> 22
<211> 239
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

<400> 22

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1   5   10   15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20   25   30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
            35   40   45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
        50   55   60

Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
65   70   75   80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85   90   95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
                100   105   110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115   120   125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130   135   140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145   150   155   160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165   170   175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly

50

```
            180                   185                   190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
        195                 200                 205


Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
        210                 215                 220


Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235


<210>  23
<211>  7010
<212>  DNA
<213>  Artificial

<220>
<223>  Engineered sequence form virus and plasmid


<220>
<221>  CDS
<222>  (206)..(1066)
<223>  Ampicillin resistance

<400>  23
ttgaagacga aagggcctcg tgatacgcct attttttatag gttaatgtca tgataataat      60

ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt     120

attttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct     180

tcaataatat tgaaaaagga agagt atg agt att caa cat ttc cgt gtc gcc     232
                           Met Ser Ile Gln His Phe Arg Val Ala
                            1               5

ctt att ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac cca     280
Leu Ile Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro
10                  15                  20                  25

gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gca cga     328
Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg
                30                  35                  40

gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag agt     376
Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser
            45                  50                  55

ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg     424
Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu
        60                  65                  70

cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg caa gag caa ctc     472
```

51

```
Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly Gln Glu Gln Leu
    75                  80                  85

ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag tac tca cca    520
Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro
90                  95                  100                 105

gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta tgc    568
Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys
                110                 115                 120

agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg    616
Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu
                125                 130                 135

aca acg atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg    664
Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met
                140                 145                 150

ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa    712
Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu
        155                 160                 165

gcc ata cca aac gac gag cgt gac acc acg atg cct gca gca atg gca    760
Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Ala Ala Met Ala
170                 175                 180                 185

aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc    808
Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser
                190                 195                 200

cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca    856
Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro
                205                 210                 215

ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct    904
Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser
            220                 225                 230

gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca    952
Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro
        235                 240                 245

gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag    1000
Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln
250                 255                 260                 265

gca act atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca    1048
Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser
                270                 275                 280

ctg att aag cat tgg taa ctgtcagacc aagtttactc atatatactt    1096
Leu Ile Lys His Trp
                285


tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat    1156
```

```
aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta   1216

gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa   1276

acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt   1336

tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag   1396

ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta   1456

atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca   1516

agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag   1576

cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa   1636

agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga   1696

acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta gtcctgtc     1756

gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc   1816

ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt   1876

gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt   1936

gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag   1996

gaagcggaag agcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac   2056

cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagtataca   2116

ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgcca acacccgctg   2176

acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct   2236

ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg aggcagccag   2296

cttacctccc ggtggtgggt cggtggtccc tgggcagggg tctcccgatc ccggacgagc   2356

ccccaaatga aagacccccg ctgacgggta gtcaatcact cagaggagac cctcccaagg   2416

aacagcgaga ccacaagtcg gatgcaactg caagagggtt tattggatac acgggtaccc   2476

gggcgactca gtcaatcgga ggactggcgc cccgagtgag gggttgtggg ctcttttatt   2536

gagctcgggg agcagaagcg cgcgaacaga agcgagaagc gaactgattg gttagttcaa   2596

ataaggcaca gggtcatttc aggtccttgg ggcaccctgg aaacatctga tggttctcta   2656

gaaactgctg agggctggac cgcatctggg gaccatctgt tcttggccct gagccggggc   2716

aggaactgct taccacagat atcctgtttg gcccatattc agctgttcca tctgttcttg   2776

gccctgagcc ggggcaggaa ctgcttacca cagatatcct gtttggccca tattcagctg   2836

ttccatctgt tcctgacctt gatctgaact tctctattct cagttatgta tttttccatg   2896
```

53

```
ccttgcaaaa tggcgttact taagctagct tgccaaacct acaggtgggg tctttcattc   2956

cccccttttt ctggagacta aataaaatct tttattttat cgatgttatc accactttgt   3016

acaagaaagc tgggtctaga tatctcgagt gcggccgctt tacttgtaca gctcgtccat   3076

gccgagagtg atcccggcgg cggtcacgaa ctccagcagg accatgtgat cgcgcttctc   3136

gttggggtct ttgctcaggg cggactggta gctcaggtag tggttgtcgg gcagcagcac   3196

ggggccgtcg ccgatggggg tgttctgctg gtagtggtcg gcgagctgca cgctgccgtc   3256

ctcgatgttg tggcggatct tgaagttcac cttgatgccg ttcttctgct tgtcggccat   3316

gatatagacg ttgtggctgt tgtagttgta ctccagcttg tgccccagga tgttgccgtc   3376

ctccttgaag tcgatgccct tcagctcgat gcggttcacc agggtgtcgc cctcgaactt   3436

cacctcggcg cgggtcttgt agttgccgtc gtccttgaag aagatggtgc gctcctggac   3496

gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc ttcatgtggt cggggtagcg   3556

ggcgaagcac tgcaggccgt agccgaaggt ggtcacgagg gtgggccagg cacgggcag    3616

cttgccggtg gtgcagatga acttcagggt cagcttgccg taggtggcat cgccctcgcc   3676

ctcgccggac acgctgaact tgtggccgtt tacgtcgccg tccagctcga ccaggatggg   3736

caccaccccg gtgaacagct cctcgccctt gctcaccatg gttaagtatt tccttagaga   3796

attggttcta tctccttcga agcctgcttt tttgtacaaa cttgtgataa caagcttctc   3856

gaggcaggac cgagagctga gcacgagggt ctcaggagaa gaacgcttct ggtgccgcta   3916

actcgacagg ccggcattt tattaattta tcagcaggtg aggtcagcag ctcagggtgt    3976

ggatttttcg ggaaagggga agttggggag gtacggggat tctagccacc agctgccgtc   4036

accagctttt ctggcagctg acgtctctgt ctggtttacg gggcggtggc ggcgcctagg   4096

tcggcatgat ctttcataca aatgattccc cctcaaaaaa ggcgggagag ccattcattt   4156

tctagcagtt gtaattttaa gggggaatca agtcttattt ttctgctagg atccgcctcg   4216

accaaagcgg ccatcgtgcc tccccactcc tgcagttcgg gggcatggat gcgcggatag   4276

ccgctgctgg tttcctggat gccgacggat ttgcactgcc ggtagaactc cgcgaggtcg   4336

tccagcctca ggcagcagct gaaccaactc gcgaggggat cgagcccggg gtgggcgaag   4396

aactccagca tgagatcccc gcgctggagg atcatccagc cggcgtcccg gaaaacgatt   4456

ccgaagccca acctttcata gaaggcggcg gtggaatcga aatctcgtga tggcaggttg   4516

ggcgtcgctt ggtcggtcat ttcgaacccc agagtcccgc tcagaagaac tcgtcaagaa   4576
```

54

```
ggcgatagaa ggcgatgcgc tgcgaatcgg gagcggcgat accgtaaagc acgaggaagc    4636

ggtcagccca ttcgccgcca agctcttcag caatatcacg ggtagccaac gctatgtcct    4696

gatagcggtc cgccacaccc agccggccac agtcgatgaa tccagaaaag cggccatttt    4756

ccaccatgat attcggcaag caggcatcgc catgggtcac gacgagatcc tcgccgtcgg    4816

gcatgcgcgc cttgagcctg gcgaacagtt cggctggcgc gagcccctga tgctcttcgt    4876

ccagatcatc ctgatcgaca agaccggctt ccatccgagt acgtgctcgc tcgatgcgat    4936

gtttcgcttg gtggtcgaat gggcaggtag ccggatcaag cgtatgcagc cgccgcattg    4996

catcagccat gatggatact ttctcggcag gagcaaggtg agatgacagg agatcctgcc    5056

ccggcacttc gcccaatagc agccagtccc ttcccgcttc agtgacaacg tcgagcacag    5116

ctgcgcaagg aacgcccgtc gtggccagcc acgatagccg cgctgcctcg tcctgcagtt    5176

cattcagggc accggacagg tcggtcttga caaaaagaac cgggcgcccc tgcgctgaca    5236

gccggaacac ggcggcatca gagcagccga ttgtctgttg tgcccagtca tagccgaata    5296

gcctctccac ccaagcggcc ggagaacctg cgtgcaatcc atcttgttca atcatgcgaa    5356

acgatcctca tcctgtctct tgatcagatc ggaattccgg cgcctagaga aggagtgagg    5416

gctggataaa gggaggatcg aggcggggtc gaacgaggag gttcaagggg gagagacggg    5476

gcggatggag gaagaggagg cggaggctta gggtgtacaa agggcttgac ccagggaggg    5536

gggtcaaaag ccaaggcttc ccaggtcacg atgtagggga cctggtctgg gtgtccatgc    5596

gggccaggtg aaaagacctt gatcttaacc tgggtgatga ggtctcggtt aaaggtgccg    5656

tctcgcggcc atccgacgtt aaaggttggc cattctgcag agcagaaggt aacccaacgt    5716

ctcttcttga catctaccga ctggttgtga gcgatccgct cgacatcttt ccagtgacct    5776

aaggtcaaac ttaagggagt ggtaacagtc tggccctaat tttcagacaa atacagaaac    5836

acagtcagac agagacaaca cagaacgatg ctgcagcgct gcagcagaca agacgcgcgg    5896

cttcggttcc aaaccgaaag caaaaattca dacggaggcg ggaactgttt taggttctcg    5956

tctcctacca gaaccacata tcctgacggg gtcggattcc acatcgactc ccttcctcag    6016

gtcgggccac aaaaacggcc cccaaagtcc ctgggacgtc tcccagggtt gcggccgggt    6076

gttcagaact cgtcagttcc accacgggtc cgccagatac agagctagtt agctaactag    6136

tacagacgca ggcgcataac atcaaacata gacactagac aatcggacag acacagataa    6196

gttgctggcc agcttacctc ccggtggtgg gtcggtggtc cctgggcagg ggtctccaaa    6256

tcccggacga gcccccaaat gaaagacccc cgtcgtgggt agtcaatcac tcagaggaga    6316
```

55

```
ccctcccaag gaacagcgag accacgattc ggatgcaaac agcaagaggc tttattggga    6376

atacgggtac ccgggcgacg cagtctatcg gaagactggc gcgccgagtg aggggttgtg    6436

ggctctttta ttgagctcgg agagcggaag cgcgcgaaca gaagcgagaa gcgaactgat    6496

tggttagttc aaataaggta cagggtcatt ttcaggtcct tggggcaccc tggaaacatc    6556

tgatgattca ctagaaactg ctgagggctg gaccgcatct ggggaccatc tgttcttggc    6616

cccgagccgg ggcaggaact gcttaccaca gatatcctgt ttggcccatc actcagctgt    6676

ctcatctgtt cttggccctg agccgggca ggaaccgctt accacagata tcctgtttgg    6736

tattcagctg tttctttgtt cctgaccttg atctgaactt ttctattctc agttatgtat    6796

ttttccatgc cttgcaaagt ggcgttactt aagctagctt gccacctacg ggtggggtct    6856

ttcaaaagca agaaacggaa agggccgcgg ctttggctcc ggtgagtgtg gggaataggg    6916

tagagtggtc taagaggcag aagcccgcga atttgggagt gtgaggggga cacatttgga    6976

ggacagtttt cggtgatctg gtatgaatta attc                                7010
```

<210> 24
<211> 286
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

<400> 24

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5                   10                  15


Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30


Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45


Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60


Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80


Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
```

56

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100   105   110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
            115   120   125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
        130   135   140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145   150   155   160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
            165   170   175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180   185   190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195   200   205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210   215   220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225   230   235   240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
            245   250   255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260   265   270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275   280   285

<210>   25
<211>   7121
<212>   DNA
<213>   Artificial

<220>

57

<223>   Engineered sequence from virus and plasmid

<220>
<221>   LTR
<222>   (149)..(737)
<223>   5' MoMuSVLTR

<220>
<221>   CDS
<222>   (1660)..(2454)
<223>   Neomycin resistance

<220>
<221>   promoter
<222>   (2806)..(3261)
<223>   HIV promoter

<220>
<221>   misc_feature
<222>   (3281)..(3305)
<223>   attB1

<220>
<221>   CDS
<222>   (3347)..(4066)
<223>   EYFP; enhanced yellow fluorescent protein

<220>
<221>   misc_feature
<222>   (4091)..(4115)
<223>   attB2

<220>
<221>   LTR
<222>   (4167)..(4760)
<223>   3' MoMuLVLTR

<400>   25
gaattaattc ataccagatc accgaaaact gtcctccaaa tgtgtccccc tcacactccc   60

aaattcgcgg gcttctgcct cttagaccac tctaccctat tccccacact caccggagcc   120

aaagccgcgg cccttccgtt tctttgcttt tgaaagaccc cacccgtagg tggcaagcta   180

gcttaagtaa cgccactttg caaggcatgg aaaaatacat aactgagaat agaaaagttc   240

agatcaaggt caggaacaaa gaaacagctg ataccaaac aggatatctg tggtaagcgg   300

ttcctgcccc ggctcagggc caagaacaga tgagacagct gagtgatggg ccaaacagga   360

tatctgtggt aagcagttcc tgccccggct cggggccaag aacagatggt ccccagatgc   420

ggtccagccc tcagcagttt ctagtgaatc atcagatgtt ccagggtgc cccaaggacc   480

tgaaaatgac cctgtacctt atttgaacta accaatcagt tcgcttctcg cttctgttcg   540

58

```
cgcgcttccg ctctccgagc tcaataaaag agcccacaac ccctcactcg gcgcgccagt    600

cttccgatag actgcgtcgc ccgggtaccc gtattcccaa taaagcctct tgctgtttgc    660

atccgaatcg tggtctcgct gttccttggg agggtctcct ctgagtgatt gactacccac    720

gacggggggtc tttcatttgg gggctcgtcc gggatttgga gaccccctgcc cagggaccac    780

cgacccacca ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag    840

tgtctatgtt tgatgttatg cgcctgcgtc tgtactagtt agctaactag ctctgtatct    900

ggcggacccg tggtggaact gacgagttct gaacacccgg ccgcaaccct gggagacgtc    960

ccagggactt tggggggccgt ttttgtggcc cgacctgagg aagggagtcg atgtggaatc   1020

cgaccccgtc aggatatgtg gttctggtag gagacgagaa cctaaaacag ttcccgcctc   1080

cgtctgaatt tttgctttcg gtttggaacc gaagccgcgc gtcttgtctg ctgcagcgct   1140

gcagcatcgt tctgtgttgt ctctgtctga ctgtgtttct gtatttgtct gaaaattagg   1200

gccagactgt taccactccc ttaagtttga ccttaggtca ctggaaagat gtcgagcgga   1260

tcgctcacaa ccagtcggta gatgtcaaga agagacgttg ggttaccttc tgctctgcag   1320

aatggccaac ctttaacgtc ggatggccgc gagacggcac ctttaaccga gacctcatca   1380

cccaggttaa gatcaaggtc ttttcacctg gcccgcatgg acacccagac caggtcccct   1440

acatcgtgac ctgggaagcc ttggcttttg acccccctcc ctgggtcaag ccctttgtac   1500

accctaagcc tccgcctcct cttcctccat ccgccccgtc tctcccccctt gaacctcctc   1560

gttcgacccc gcctcgatcc tccctttatc cagccctcac tccttctcta ggcgccggaa   1620

ttccgatctg atcaagagac aggatgagga tcgtttcgc atg att gaa caa gat    1674
                                             Met Ile Glu Gln Asp
                                              1               5

gga ttg cac gca ggt tct ccg gcc gct tgg gtg gag agg cta ttc ggc    1722
Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly
             10              15                  20

tat gac tgg gca caa cag aca atc ggc tgc tct gat gcc gcc gtg ttc    1770
Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe
             25              30                  35

cgg ctg tca gcg cag ggg cgc ccg gtt ctt ttt gtc aag acc gac ctg    1818
Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu
         40              45              50

tcc ggt gcc ctg aat gaa ctg cag gac gag gca gcg cgg cta tcg tgg    1866
Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp
 55              60              65

ctg gcc acg acg ggc gtt cct tgc gca gct gtg ctc gac gtt gtc act    1914


                                    59
```

```
Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val Thr
70              75                  80                  85

gaa gcg gga agg gac tgg ctg cta ttg ggc gaa gtg ccg ggg cag gat        1962
Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu Val Pro Gly Gln Asp
                    90                  95                  100

ctc ctg tca tct cac ctt gct cct gcc gag aaa gta tcc atc atg gct        2010
Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met Ala
                105             110             115

gat gca atg cgg cgg ctg cat acg ctt gat ccg gct acc tgc cca ttc        2058
Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro Phe
            120             125             130

gac cac caa gcg aaa cat cgc atc gag cga gca cgt act cgg atg gaa        2106
Asp His Gln Ala Lys His Arg Ile Glu Arg Ala Arg Thr Arg Met Glu
        135             140             145

gcc ggt ctt gtc gat cag gat gat ctg gac gaa gag cat cag ggg ctc        2154
Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly Leu
150             155             160             165

gcg cca gcc gaa ctg ttc gcc agg ctc aag gcg cgc atg ccc gac ggc        2202
Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp Gly
                170             175             180

gag gat ctc gtc gtg acc cat ggc gat gcc tgc ttg ccg aat atc atg        2250
Glu Asp Leu Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile Met
            185             190             195

gtg gaa aat ggc cgc ttt tct gga ttc atc gac tgt ggc cgg ctg ggt        2298
Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly
        200             205             210

gtg gcg gac cgc tat cag gac ata gcg ttg gct acc cgt gat att gct        2346
Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile Ala
    215             220             225

gaa gag ctt ggc ggc gaa tgg gct gac cgc ttc ctc gtg ctt tac ggt        2394
Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr Gly
230             235             240             245

atc gcc gct ccc gat tcg cag cgc atc gcc ttc tat cgc ctt ctt gac        2442
Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp
            250             255             260

gag ttc ttc tga gcgggactct ggggttcgaa atgaccgacc aagcgacgcc           2494
Glu Phe Phe


caacctgcca tcacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg       2554

aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt       2614

cttcgcccac cccgggctcg atccctcgc gagttggttc agctgctgcc tgaggctgga        2674
```

cgacctcgcg gagttctacc ggcagtgcaa atccgtcggc atccaggaaa ccagcagcgg   2734

ctatccgcgc atccatgccc ccgaactgca ggagtgggga ggcacgatgg ccgctttggt   2794

cgaggcggat cctggaaggg ctaatttggt cccaaagaag acaagagatc cttgatctgt   2854

ggatctacca cacacaaggc tacttccctg attggcagaa ttacacacca gggccaggga   2914

tcagatatcc actgaccttt ggatggtgct tcaagctagt accagttgag ccagagaagg   2974

tagaagaggc caatgaagga gagaacaaca gcttgttaca ccctatgagc ctgcatggga   3034

tggaggacgc ggagaaagaa gtgttagtgt ggaggtttga cagcaaacta gcatttcatc   3094

acatggcccg agagctgcat ccggagtact acaaagactg ctgacatcga gctttctaca   3154

agggactttc cgctggggac tttccaggga ggcgtggcct gggcgggact ggggagtggc   3214

gtccctcaga tgctgcatat aagcagctgc tttttgcctg tactgggcct cgagaagctt   3274

gttatcacaa gtttgtacaa aaaagcaggc ttcgaaggag atagaaccaa ttctctaagg   3334

aaatacttaa cc atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg   3385
             Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val
             265   270   275

ccc atc ctg gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc   3433
Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser
    280   285   290

gtg tcc ggc gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg   3481
Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu
    295   300   305

aag ttc atc tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc   3529
Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu
310   315   320   325

gtg acc acc ttc ggc tac ggc ctg cag tgc ttc gcc cgc tac ccc gac   3577
Val Thr Thr Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp
                330   335   340

cac atg aag cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac   3625
His Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr
                345   350   355

gtc cag gag cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc   3673
Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr
        360   365   370

cgc gcc gag gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag   3721
Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu
    375   380   385

ctg aag ggc atc gac ttc aag gag gac ggc aac atc ctg ggg cac aag   3769
Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys
390   395   400   405

61

```
ctg gag tac aac tac aac agc cac aac gtc tat atc atg gcc gac aag      3817
Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys
                410             415             420

cag aag aac ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag      3865
Gln Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu
                425             430             435

gac ggc agc gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc      3913
Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile
            440             445             450

ggc gac ggc ccc gtg ctg ctg ccc gac aac cac tac ctg agc tac cag      3961
Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln
        455             460             465

tcc gcc ctg agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg      4009
Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu
470             475             480             485

ctg gag ttc gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg      4057
Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu
                490             495             500

tac aag taa agcggccgca ctcgagatat ctagacccag ctttcttgta             4106
Tyr Lys


caaagtggtg ataacatcga taaaataaaa gattttattt agtctccaga aaaagggggg   4166

aatgaaagac cccacctgta ggtttggcaa gctagcttaa gtaacgccat tttgcaaggc   4226

atggaaaaat acataactga gaatagagaa gttcagatca aggtcaggaa cagatggaac   4286

agctgaatat gggccaaaca ggatatctgt ggtaagcagt tcctgccccg gctcagggcc   4346

aagaacagat ggaacagctg aatatgggcc aaacaggata tctgtggtaa gcagttcctg   4406

ccccggctca gggccaagaa cagatggtcc ccagatgcgg tccagccctc agcagtttct   4466

agagaaccat cagatgtttc cagggtgccc caaggacctg aaatgaccct gtgccttatt   4526

tgaactaacc aatcagttcg cttctcgctt ctgttcgcgc gcttctgctc cccgagctca   4586

ataaaagagc ccacaacccc tcactcgggg cgccagtcct ccgattgact gagtcgcccg   4646

ggtacccgtg tatccaataa accctcttgc agttgcatcc gacttgtggt ctcgctgttc   4706

cttgggaggg tctcctctga gtgattgact acccgtcagc gggggtcttt catttggggg   4766

ctcgtccggg atcgggagac ccctgcccag gaccaccga cccaccaccg ggaggtaagc     4826

tggctgcctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc agctcccgga   4886

gacggtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc agggcgcgtc   4946
```

62

```
agcgggtgtt ggcgggtgtc ggggcgcagc catgacccag tcacgtagcg atagcggagt     5006

gtatactggc ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg     5066

tgtgaaatac cgcacagatg cgtaaggaga aaataccgca tcaggcgctc ttccgcttcc     5126

tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc agctcactca     5186

aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca     5246

aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg     5306

ctccgccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg     5366

acaggactat aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt     5426

ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt     5486

tctcatagct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc     5546

tgtgtgcacg aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt     5606

gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt     5666

agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc     5726

tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa     5786

agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt     5846

tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct     5906

acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta     5966

tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa     6026

agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc     6086

tcagcgatct gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact     6146

acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc     6206

tcaccggctc cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt     6266

ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta     6326

agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctgcagg catcgtggtg     6386

tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt     6446

acatgatccc ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc     6506

agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt     6566

actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc     6626

tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaacacg ggataatacc     6686
```

63

```
gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa    6746

ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac    6806

tgatcttcag catctttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa    6866

aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt    6926

tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa    6986

tgtatttaga aaaataaaca aatagggggtt ccgcgcacat ttccccgaaa agtgccacct    7046

gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg    7106

cccctttcgtc ttcaa                                                    7121
```

<210> 26
<211> 264
<212> PRT
<213> Artificial

<220>
<223> Synthetic Construct

<400> 26

```
Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val
1               5                   10                  15


Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
            20                  25                  30


Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
        35                  40                  45


Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
    50                  55                  60


Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val
65                  70                  75                  80


Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
                85                  90                  95


Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
                100                 105                 110


Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
```
```

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
    130   135   140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu
145   150   155   160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
                165   170   175

Arg Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
            180   185   190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
        195   200   205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
    210   215   220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe
225   230   235   240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
            245   250   255

Tyr Arg Leu Leu Asp Glu Phe Phe
            260

<210>   27
<211>   239
<212>   PRT
<213>   Artificial

<220>
<223>   Synthetic Construct

<400>   27

Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1   5   10   15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20   25   30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
    35   40   45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50   55   60

Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
65   70   75   80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
           85   90   95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
        100   105   110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
      115   120   125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
    130   135   140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145   150   155   160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
         165   170   175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
         180   185   190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
      195   200   205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210   215   220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225   230   235

<210> 28
<211> 7121
<212> DNA
<213> Artificial

<220>
<223> Engineered sequence form virus and plasmid

<220>
<221> CDS
<222> (206)..(1066)
<223> Ampicillin resistance

<400> 28
```
ttgaagacga aagggcctcg tgatacgcct atttttatag gttaatgtca tgataataat      60

ggtttcttag acgtcaggtg cacttttcg gggaaatgtg cgcggaaccc ctatttgttt      120

atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct      180

tcaataatat tgaaaaagga agagt atg agt att caa cat ttc cgt gtc gcc      232
                           Met Ser Ile Gln His Phe Arg Val Ala
                           1               5

ctt att ccc ttt ttt gcg gca ttt tgc ctt cct gtt ttt gct cac cca      280
Leu Ile Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro
10              15                  20                  25

gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg ggt gca cga      328
Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg
            30                  35                  40

gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc ctt gag agt      376
Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser
                45                  50                  55

ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt aaa gtt ctg      424
Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu
            60                  65                  70

cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg caa gag caa ctc      472
Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly Gln Glu Gln Leu
    75                  80                  85

ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag tac tca cca      520
Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro
90                  95                  100                 105

gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga gaa tta tgc      568
Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys
                110                 115                 120

agt gct gcc ata acc atg agt gat aac act gcg gcc aac tta ctt ctg      616
Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu
            125                 130                 135

aca acg atc gga gga ccg aag gag cta acc gct ttt ttg cac aac atg      664
Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met
        140                 145                 150

ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg gag ctg aat gaa      712
```

67

```
Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu
    155                 160             165

gcc ata cca aac gac gag cgt gac acc acg atg cct gca gca atg gca       760
Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Ala Ala Met Ala
170                 175             180                 185

aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act cta gct tcc       808
Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser
                190             195                 200

cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt gca gga cca       856
Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro
                205             210                 215

ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct gat aaa tct       904
Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser
            220             225             230

gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca ctg ggg cca       952
Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro
    235             240             245

gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg ggg agt cag      1000
Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln
250                 255             260                 265

gca act atg gat gaa cga aat aga cag atc gct gag ata ggt gcc tca      1048
Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser
                270             275                 280

ctg att aag cat tgg taa ctgtcagacc aagtttactc atatatactt            1096
Leu Ile Lys His Trp
                285

tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat cctttttgat   1156

aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc agaccccgta   1216

gaaaagatca aaggatcttc ttgagatcct tttttctgc gcgtaatctg ctgcttgcaa    1276

acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct accaactctt    1336

tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct tctagtgtag    1396

ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct cgctctgcta    1456

atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg gttggactca    1516

agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc gtgcacacag    1576

cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga gctatgagaa    1636

agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga    1696

acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc    1756
```

68

```
gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc   1816

ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg ctggcctttt   1876

gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat taccgccttt   1936

gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc agtgagcgag   1996

gaagcggaag agcgcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac   2056

cgcatatggt gcactctcag tacaatctgc tctgatgccg catagttaag ccagtataca   2116

ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgcca cacccgctg    2176

acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct   2236

ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg aggcagccag   2296

cttacctccc ggtggtgggt cggtggtccc tgggcagggg tctcccgatc ccggacgagc   2356

ccccaaatga agacccccg ctgacgggta gtcaatcact cagaggagac cctcccaagg    2416

aacagcgaga ccacaagtcg gatgcaactg caagagggtt tattggatac acgggtaccc   2476

gggcgactca gtcaatcgga ggactggcgc cccgagtgag gggttgtggg ctcttttatt   2536

gagctcgggg agcagaagcg cgcgaacaga agcgagaagc gaactgattg gttagttcaa   2596

ataaggcaca gggtcatttc aggtccttgg ggcaccctgg aaacatctga tggttctcta   2656

gaaactgctg agggctggac cgcatctggg gaccatctgt tcttggccct gagccggggc   2716

aggaactgct taccacagat atcctgtttg gcccatattc agctgttcca tctgttcttg   2776

gccctgagcc ggggcaggaa ctgcttacca cagatatcct gtttggccca tattcagctg   2836

ttccatctgt tcctgacctt gatctgaact tctctattct cagttatgta tttttccatg   2896

ccttgcaaaa tggcgttact taagctagct tgccaaacct acaggtgggg tctttcattc   2956

ccccctttt ctggagacta aataaaatct tttatttat cgatgttatc accactttgt     3016

acaagaaagc tgggtctaga tatctcgagt gcggccgctt tacttgtaca gctcgtccat   3076

gccgagagtg atcccggcgg cggtcacgaa ctccagcagg accatgtgat cgcgcttctc   3136

gttggggtct ttgctcaggg cggactggta gctcaggtag tggttgtcgg gcagcagcac   3196

ggggccgtcg ccgatggggg tgttctgctg gtagtggtcg gcgagctgca cgctgccgtc   3256

ctcgatgttg tggcggatct tgaagttcac cttgatgccg ttcttctgct tgtcggccat   3316

gatatagacg ttgtggctgt tgtagttgta ctccagcttg tgccccagga tgttgccgtc   3376

ctccttgaag tcgatgccct tcagctcgat gcggttcacc agggtgtcgc cctcgaactt   3436

cacctcggcg cgggtcttgt agttgccgtc gtccttgaag aagatggtgc gctcctggac   3496
```

69

```
gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc ttcatgtggt cggggtagcg   3556

ggcgaagcac tgcaggccgt agccgaaggt ggtcacgagg gtgggccagg gcacgggcag   3616

cttgccggtg gtgcagatga acttcagggt cagcttgccg taggtggcat cgccctcgcc   3676

ctcgccggac acgctgaact tgtggccgtt tacgtcgccg tccagctcga ccaggatggg   3736

caccaccccg gtgaacagct cctcgccctt gctcaccatg gttaagtatt tccttagaga   3796

attggttcta tctccttcga agcctgcttt tttgtacaaa cttgtgataa caagcttctc   3856

gaggcccagt acaggcaaaa agcagctgct tatatgcagc atctgaggga cgccactccc   3916

cagtcccgcc caggccacgc ctccctggaa agtccccagc ggaaagtccc ttgtagaaag   3976

ctcgatgtca gcagtctttg tagtactccg gatgcagctc tcgggccatg tgatgaaatg   4036

ctagtttgct gtcaaacctc cacactaaca cttctttctc cgcgtcctcc atcccatgca   4096

ggctcatagg gtgtaacaag ctgttgttct ctccttcatt ggcctcttct accttctctg   4156

gctcaactgg tactagcttg aagcaccatc caaaggtcag tggatatctg atccctggcc   4216

ctggtgtgta attctgccaa tcagggaagt agccttgtgt gtggtagatc cacagatcaa   4276

ggatctcttg tcttctttgg gaccaaatta gcccttccag gatccgcctc gaccaaagcg   4336

gccatcgtgc ctccccactc ctgcagttcg ggggcatgga tgcgcggata gccgctgctg   4396

gtttcctgga tgccgacgga tttgcactgc cggtagaact ccgcgaggtc gtccagcctc   4456

aggcagcagc tgaaccaact cgcgagggga tcgagcccgg ggtgggcgaa gaactccagc   4516

atgagatccc cgcgctggag gatcatccag ccggcgtccc ggaaaacgat tccgaagccc   4576

aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcaggtt gggcgtcgct   4636

tggtcggtca tttcgaaccc cagagtcccg ctcagaagaa ctcgtcaaga aggcgataga   4696

aggcgatgcg ctgcgaatcg ggagcggcga taccgtaaag cacgaggaag cggtcagccc   4756

attcgccgcc aagctcttca gcaatatcac gggtagccaa cgctatgtcc tgatagcggt   4816

ccgccacacc cagccggcca cagtcgatga atccagaaaa gcggccattt tccaccatga   4876

tattcggcaa gcaggcatcg ccatgggtca cgacgagatc ctcgccgtcg ggcatgcgcg   4936

ccttgagcct ggcgaacagt tcggctggcg cgagcccctg atgctcttcg tccagatcat   4996

cctgatcgac aagaccggct tccatccgag tacgtgctcg ctcgatgcga tgtttcgctt   5056

ggtggtcgaa tgggcaggta gccggatcaa gcgtatgcag ccgccgcatt gcatcagcca   5116

tgatggatac tttctcggca ggagcaaggt gagatgacag gagatcctgc cccggcactt   5176
```

70

```
cgcccaatag cagccagtcc cttcccgctt cagtgacaac gtcgagcaca gctgcgcaag     5236

gaacgcccgt cgtggccagc cacgatagcc gcgctgcctc gtcctgcagt tcattcaggg     5296

caccggacag gtcggtcttg acaaaaagaa ccgggcgccc ctgcgctgac agccggaaca     5356

cggcggcatc agagcagccg attgtctgtt gtgcccagtc atagccgaat agcctctcca     5416

cccaagcggc cggagaacct gcgtgcaatc catcttgttc aatcatgcga aacgatcctc     5476

atcctgtctc ttgatcagat cggaattccg gcgcctagag aaggagtgag ggctggataa     5536

agggaggatc gaggcggggt cgaacgagga ggttcaaggg ggagagacgg ggcggatgga     5596

ggaagaggag gcggaggctt agggtgtaca aagggcttga cccagggagg ggggtcaaaa     5656

gccaaggctt cccaggtcac gatgtagggg acctggtctg ggtgtccatg cgggccaggt     5716

gaaaagacct tgatcttaac ctgggtgatg aggtctcggt taaaggtgcc gtctcgcggc     5776

catccgacgt taaaggttgg ccattctgca gagcagaagg taacccaacg tctcttcttg ·   5836

acatctaccg actggttgtg agcgatccgc tcgacatctt tccagtgacc taaggtcaaa     5896

cttaagggag tggtaacagt ctggccctaa ttttcagaca aatacagaaa cacagtcaga     5956

cagagacaac acagaacgat gctgcagcgc tgcagcagac aagacgcgcg gcttcggttc     6016

caaaccgaaa gcaaaaattc agacggaggc gggaactgtt ttaggttctc gtctcctacc     6076

agaaccacat atcctgacgg ggtcggattc cacatcgact cccttcctca ggtcgggcca     6136

caaaaacggc ccccaaagtc cctgggacgt ctcccagggt tgcggccggg tgttcagaac     6196

tcgtcagttc caccacgggt ccgccagata cagagctagt tagctaacta gtacagacgc     6256

aggcgcataa catcaaacat agacactaga caatcggaca gacacagata agttgctggc     6316

cagcttacct cccggtggtg ggtcggtggt ccctgggcag gggtctccaa atcccggacg     6376

agcccccaaa tgaaagaccc ccgtcgtggg tagtcaatca ctcagaggag accctcccaa     6436

ggaacagcga gaccacgatt cggatgcaaa cagcaagagg ctttattggg aatacgggta     6496

cccgggcgac gcagtctatc ggaagactgg cgcgccgagt gagggggttgt gggctctttt    6556

attgagctcg gagagcggaa gcgcgcgaac agaagcgaga agcgaactga ttggttagtt     6616

caaataaggt acagggtcat tttcaggtcc ttggggcacc ctggaaacat ctgatgattc     6676

actagaaact gctgagggct ggaccgcatc tggggaccat ctgttcttgg ccccgagccg     6736

gggcaggaac tgcttaccac agatatcctg tttggcccat cactcagctg tctcatctgt     6796

tcttggccct gagccggggc aggaaccgct taccacagat atcctgtttg gtattcagct     6856

gtttctttgt tcctgacctt gatctgaact tttctattct cagttatgta tttttccatg     6916
```

71

ccttgcaaag tggcgttact taagctagct tgccacctac gggtggggtc tttcaaaagc   6976

aaagaaacgg aagggccgcg gctttggctc cggtgagtgt ggggaatagg gtagagtggt   7036

ctaagaggca gaagcccgcg aatttgggag tgtgaggggg acacatttgg aggacagttt   7096

tcggtgatct ggtatgaatt aattc   7121

<210>   29
<211>   286
<212>   PRT
<213>   Artificial

<220>
<223>   Synthetic Construct

<400>   29

Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1   5   10   15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20   25   30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35   40   45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50   55   60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65   70   75   80

Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85   90   95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100   105   110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115   120   125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130   135   140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu

72

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
            165   170   175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180   185   190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
            195   200   205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210   215   220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225   230   235   240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
            245   250   255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260   265   270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
    275   280   285

73